

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>Publiker</u>	NA Sequence (#) <u>10</u>	STN _____
Searcher Phone #: <u>308-4740</u>	AA Sequence (#) _____	Dialog _____
Searcher Location: <u>1ED7</u>	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>6-19</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>6-20</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: <u>15</u>	Fulltext _____	Sequence Systems <u>05</u>
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: <u>15</u>	Other _____	Other (specify) _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sv model

Run on: June 19, 2001, 17:00:40 ; Search time 7917.7 Seconds
(Without alignments)
1550.862 Million cell updates/sec

Title: US-09-394-519-105
Perfect score: 1299
Sequence: 1 tcttgcacaaataaaaaa.....cagttgttagtaccctaaa 12991

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estin6:*
69: em_estin7:*
70: em_estin8:*
71: em_estin9:*
72: em_estin10:*
73: em_estin11:*
74: em_estin12:*
75: em_estin13:*
76: em_estin14:*
77: em_estin15:*
78: em_estin16:*
79: em_estin17:*
80: em_estin18:*
81: em_estin19:*
82: em_estin20:*
83: em_estin21:*
84: em_estin22:*
85: em_estin23:*
86: em_estin24:*
87: em_estin25:*
88: em_estin26:*
89: em_estin27:*
90: em_estin28:*
91: em_estin29:*
92: em_estin30:*
93: em_estin31:*
94: em_estin32:*
95: em_estin33:*
96: em_estin34:*
97: em_estin35:*
98: em_estin36:*
99: em_estin37:*
100: em_estin38:*
101: em_estin39:*
102: em_estin40:*
103: em_estin41:*
104: em_estin42:*
105: em_estin43:*
106: em_estin44:*
107: em_estin45:*
108: em_estin46:*
109: em_estin47:*
110: em_estin48:*
111: em_estin49:*
112: em_estin50:*
113: em_estin51:*
114: em_estin52:*
115: em_estin53:*
116: em_estin54:*

```

117: qb_est148:*
118: qb_est149:*
119: qb_est150:*
120: qb_est151:*
121: qb_est152:*
122: qb_est153:*
123: qb_est154:*
124: qb_est155:*
125: qb_est156:*
126: qb_est157:*
127: qb_est158:*
128: qb_est159:*
129: qb_est160:*
130: qb_est161:*
131: qb_est162:*
132: qb_est163:*
133: qb_est164:*
134: qb_est165:*
135: qb_est166:*
136: qb_est167:*
137: qb_est168:*
138: qb_est169:*
139: qb_est170:*
140: qb_est171:*
141: qb_est172:*
142: qb_est173:*
143: qb_est174:*
144: qb_est175:*
145: qb_est176:*
146: qb_est177:*
147: qb_est178:*
148: qb_est179:*
149: qb_est180:*
150: qb_est181:*
151: qb_est182:*
152: qb_est183:*
153: qb_est184:*
154: qb_est185:*
155: qb_est186:*
156: qb_est187:*
157: qb_est188:*
158: qb_est189:*
159: qb_est190:*
160: qb_est191:*
161: qb_est192:*
162: qb_est193:*
163: qb_est194:*
164: qb_est195:*
165: qb_est196:*
166: qb_est197:*
167: qb_est198:*
168: qb_est199:*
169: qb_est200:*
170: qb_est201:*
171: qb_est202:*
172: qb_est203:*
173: qb_est204:*
174: qb_est205:*
175: qb_est206:*
176: qb_est207:*
177: qb_est208:*
178: qb_est209:*
179: qb_est210:*
180: qb_est211:*
181: qb_est212:*
182: qb_est213:*
183: qb_est214:*
184: qb_est215:*
185: qb_est216:*
186: qb_est217:*
187: qb_est218:*
188: qb_est219:*
189: qb_est220:*

```

```

190: qb_est221:*
191: qb_est222:*
192: qb_est223:*
193: qb_est224:*
194: qb_est225:*
195: qb_est226:*
196: qb_est227:*
197: qb_est228:*
198: qb_est229:*
199: qb_est230:*
200: qb_est231:*
201: qb_est232:*
202: qb_est233:*
203: qb_est234:*
204: qb_est235:*
205: qb_est236:*
206: qb_est237:*
207: qb_est238:*
208: qb_est239:*
209: qb_est240:*
210: qb_est241:*
211: qb_est242:*
212: qb_est243:*
213: qb_est244:*
214: qb_est245:*
215: qb_est246:*
216: qb_est247:*
217: qb_est248:*
218: qb_est249:*
219: qb_est250:*
220: qb_est251:*
221: qb_est252:*
222: qb_est253:*
223: qb_est254:*
224: qb_est255:*
225: qb_est256:*
226: qb_est257:*
227: qb_est258:*
228: qb_est259:*
229: qb_est260:*
230: qb_est261:*
231: qb_est262:*
232: qb_est263:*
233: qb_est264:*
234: qb_est265:*
235: qb_est266:*
236: qb_est267:*
237: qb_est268:*
238: qb_est269:*
239: qb_est270:*
240: qb_est271:*
241: qb_est272:*
242: qb_est273:*
243: qb_est274:*
244: qb_est275:*
245: qb_est276:*
246: qb_est277:*
247: qb_est278:*
248: qb_est279:*
249: qb_est280:*
250: qb_est281:*
251: qb_est282:*
252: qb_est283:*
253: qb_est284:*
254: qb_est285:*
255: qb_est286:*
256: qb_est287:*
257: qb_est288:*
258: qb_est289:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	440	33.9	493	236	AO963001	AO963001 LERGI80TF
2	338.8	26.1	363	188	T14116	T14116 2281 Lambda
3	166.6	12.8	235	236	AO963002	AO963002 LERGI80TR
4	154.4	11.9	540	31	AV527594	AV527594 AV527594
5	142.6	11.0	526	113	AM233999	AM233999 sf32h1.Y
6	142.6	11.0	558	118	AM597702	AM597702 sj49g05.Y
7	135.2	10.5	448	118	AM620835	AM620835 sj47b12.Y
8	135.8	10.5	534	137	BE611146	BE611146 sq/6d03.Y
9	134.6	10.4	587	113	AM257000	AM257000 EST1305137
10	134.6	10.4	791	155	BC586052	BC586052 EST487817
11	127.4	9.8	485	238	AE127897	AE127897 OSJRB009
12	126.2	9.7	270	31	AV549228	AV549228 AV549228
13	120.6	9.3	532	16	A1161927	A1161927 A009P620
14	119	9.2	571	103	A1894846	A1894846 EST264289
15	119	9.2	766	174	BC130539	BC130539 EST476185
16	118.8	9.1	385	118	AM622377	AM622377 EST131176
17	118.4	9.1	692	155	BC592323	BC592323 EST491001
18	117	9.0	522	21	A1490285	A1490285 EST248611
19	116.2	8.9	611	111	AM065881	AM065881 687002C04
20	115.4	8.9	574	31	AV550536	AV550536 sac35e02
21	114.8	8.8	327	153	BC405537	BC405537 sac35e02
22	113.8	8.8	475	144	BF097661	BF097661 EST415734
23	113.8	8.8	583	21	A1487752	A1487752 EST246074
24	112.6	8.7	796	151	BF622756	BF622756 HVSME000
25	112.4	8.7	512	118	AM585659	AM585659 EST317282
26	111.8	8.6	658	119	AM696439	AM696439 NF106B07
27	111.4	8.6	373	153	BC405482	BC405482 sac44a11
28	110.4	8.5	348	17	A1165295	A1165295 A080P54U
29	107.2	8.3	663	105	AL507122	AL507122 AL507122
30	107	8.2	429	30	AV409036	AV409036 AV409036
31	106.6	8.2	559	158	H77029	H77029 17460 Lambda
32	106	8.2	723	155	BC586298	BC586298 EST488064
33	105.4	8.1	522	137	BE555532	BE555532 sp90e07.Y
34	102.6	7.9	564	119	AM720051	AM720051 LjNEST13a
35	101.6	7.8	434	143	BF010257	BF010257 ss96e07.Y
36	100.6	7.7	489	110	AM031324	AM031324 EST274778
37	100.4	7.7	550	20	A1461184	A1461184 sa76b10.Y
38	100.2	7.7	420	153	BC405625	BC405625 sac45e11
39	100.2	7.7	479	145	BF203107	BF203107 WHE1759-1
40	100.2	7.7	480	145	BF202003	BF202003 WHE1759-1
41	100	7.7	499	237	AZ044889	AZ044889 A363/F.SO
42	100	7.7	514	234	AO842107	AO842107 T134402.S
43	100	7.7	588	234	AO842067	AO842067 T134353.S
44	99.6	7.7	554	136	BE500188	BE500188 WHE0980.F
45	99	7.6	404	136	BE474893	BE474893 sp69g07.Y

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	SEQUENCE	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
1	AO963001/c	AO963001 493 bp DNA	28-JAN-2000	AO963001	AO963001	GI:6790702	thale cress.	Arabidopsis thaliana
LOCUS	LERGI80TF	LERGI80TF LERGI Arabidopsis thaliana genomic clone LERGI80, DNA						
DEFINITION	sequence.							
ACCESSION	AO963001							
VERSION	AO963001.1							
KEYWORDS	GI:6790702							
SOURCE	thale cress.							
ORGANISM	Arabidopsis thaliana							
REFERENCE	Arabidopsis thaliana							
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids							
TITLE	1 (bases 1 to 493)							
	Buell, C.R., Lin, X., Pal, G., Barnstead, M., Bowman, C., Utebach, T., Feldhlyum, T., Liang, F., Creasy, T. and Fraser, C.M.							
	Genomic survey sequencing of Landsberg erecta ecotype of							

JOURNAL COMMENT

Arabidopsis thaliana and identification of sequence-based polymorphisms
Unpublished (2000)
Contact: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: ateligr.org
For additional information, see <http://www.tigr.org/cdb/at.html>
Seq primer: TF
Class: shotgun.

FEATURES

Source
Location/Qualifiers
1..493
/organism="Arabidopsis thaliana"
/strain="Landsberg erecta"
/db_xref="taxon:3702"
/clone="LERGI80"
/clone_11b="LERGI"
/note="Organ: Leaf; Vector: pUC19JK; Total genomic DNA was sheared to 0.4-0.7 Kbp before ligation."
BASE COUNT 179 a 100 c 79 g 135 t
ORIGIN

Query Match 33.9%; Score 440; DB 236; Length 493;
Best Local Similarity 100.0%; Pred. No. 3.6e-108;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	858	ccatttgagagcctcctcaactccagttcattcatcatcaagcaccatagttggtgaat	917
DB	492	CCATTGAGAGCCCTCCCAACTCCAGTTTCATCATCAAGCACCATTATGCGGAAT	433
QY	918	ggcaatcgaagaagaataatgatgatgatgatgatgatgaacatgaagaagaaga	977
DB	432	GGCAATCGAAGAGATTAATGATGACCATGATGATGATGATGATGATGATGATG	373
QY	978	gaagctaccacatcgttgggttgggttgggttgggttgggttgggttgggttgggt	1037
DB	372	GAGGCTACCAACATCTGTTGGGTTGGAGTGGATGATGATGATGATGATGATGAT	313
QY	1038	gcattcgttggttcattgacataatatttgatcattgacacccatccatgagattt	1097
DB	312	GCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	253
QY	1098	aagaaagatcattcatttccatcatttccatcatttccatcatttccatcatttcc	1157
DB	252	AAGAAATGATTTACCTTTTACCCATGTTTAAAGATGAGAGGTTAACTATATAT	193
QY	1158	tgggtgcttactatctttagctcttcttcttcttcttcttcttcttcttcttctt	1217
DB	192	TGGGTGCTTACTATTATTAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	133
QY	1218	attaacttcttggttggttggttggttggttggttggttggttggttggttggtt	1277
DB	132	ATTAACCTTGTGTTGGGTGGGATGATGATGATGATGATGATGATGATGATGAT	73
QY	1278	ctcagttgtagttaccta 1297	
DB	72	CTCAGTTGTTAGTTACCTAA 53	

RESULT 2	LOCUS	DEFINITION	SEQUENCE	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
2	T14116	T14116 363 bp mRNA	EST					
LOCUS	2281	Lambda-PR2 Arabidopsis thaliana cDNA clone 47E1077, mRNA						
DEFINITION	sequence.							
ACCESSION	T14116							
VERSION	T14116.1							
KEYWORDS	GI:931090							
SOURCE	thale cress.							
ORGANISM	Arabidopsis thaliana							
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;							
AUTHORS								
TITLE								

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 363)

REFERENCE
AUTHORS
Newman, T., deBrujn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Ralke, N., Somerville, S., Thomas, M., Retzel, E. and Somerville, C.

TITLE
JOURNAL
MEDLINE
COMMENT
95148729

On Nov 29, 1993 this sequence version replaced gi:430485.

Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRC, Michigan State University, Plant Biology Bldg., E. Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 223j3tcn@lhm.cl.msu.edu

Seq primer: 77.

FEATURES
source
1. 363
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="var columbica"
/db_xref="taxon:3702"
/clone="47E1077"
/note="Vector: lambda zip-lox; Site 1: Sal; Site 2: NotI; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRL's lambda zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA."

BASE COUNT
106 a 90 c 73 g 85 t 9 others

ORIGIN

Query Match 26.1%; Score 338.8; DB 188; Length 363;
Best Local Similarity 96.7%; Pred. No. 8e-81;
Matches 351; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

Db 1 tcttgcacaaataaataaagaaggttggtaggaacaaagaaagatcgtga 60
1 tnttgcacaaataaataaagaaggttggtaggaacaaagaaagatcgtga 60

QY 61 caaacacatgctaaacacacagtcgaatcctctctctgctactctgtgaaga 120
61 caaacacatgctaaacacacagtcgaatcctctctctgctactctgtgaaga 120

Db 61 caaacacatgctaaacacacagtcgaatcctctctctgctactctgtgaaga 120
61 caaacacatgctaaacacacagtcgaatcctctctctgctactctgtgaaga 120

QY 121 aggcgaacccggaagaagaccaggttgaactgtccaagatgaaccccttaacacca 180
121 aggcgaacccggaagaagaccaggttgaactgtccaagatgaaccccttaacacca 180

Db 121 aggcgaacccggaagaagaccaggttgaactgtccaagatgaaccccttaacacca 180
121 aggcgaacccggaagaagaccaggttgaactgtccaagatgaaccccttaacacca 180

QY 181 agttcgttactaacaactacagctgaagccagaggaacttttgaagactga 240
181 agttcgttactaacaactacagctgaagccagaggaacttttgaagactga 240

Db 181 agttcgttactaacaactacagctgaagccagaggaacttttgaagactga 240
181 agttcgttactaacaactacagctgaagccagaggaacttttgaagactga 240

QY 241 ggaagatggaacccaggtgtgtccctcagaagacaccccgctgggtggcgagtcgca 300
241 ggaagatggaacccaggtgtgtccctcagaagacaccccgctgggtggcgagtcgca 300

Db 241 ggaagatggaacccaggtgtgtccctcagaagacaccccgctgggtggcgagtcgca 300
241 ggaagatggaacccaggtgtgtccctcagaagacaccccgctgggtggcgagtcgca 300

QY 301 agaacagaagatcttcccaatctctcctctcaccctcttc-gtctctcttca 359
301 agaacagaagatcttcccaatctctcctctcaccctcttc-gtctctcttca 359

Db 301 agaacagaagatcttcccaatctctcctctcaccctcttc-gtctctcttca 359
301 agaacagaagatcttcccaatctctcctctcaccctcttc-gtctctcttca 359

QY 360 aag 362
360 aag 362

Db 361 aag 363

RESULT 3
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AO963002 255 bp DNA GSS 28-JAN-2000
LEK180TR LERG Arabidopsis thaliana genomic clone LEK180, DNA sequence.
AO963002
AO963002.1 GI:6790703
GSS

thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 255)

Buell, C.R., Lin, X., Pal, G., Barnstead, M., Bowman, C., Uterbach, T., Feldblyum, T., Liang, F., Greasy, T. and Fraser, C.M.
Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
Unpublished (2000)

Contact: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: ateflgr.org
For additional information, see <http://www.tigr.org/tdb/ac/at.html>
Seq primer: TR
Class: shotgun.

FEATURES
source
1. 255
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="Landsberg erecta"
/db_xref="taxon:3702"
/clone="LEK180"
/note="Organ: Leaf; Vector: pUC19UK; Total genomic DNA was sheared to 0.4-0.7 kbp before ligation."

BASE COUNT
60 a 61 c 58 g 75 t 1 others

ORIGIN

Query Match 12.8%; Score 166.6; DB 236; Length 255;
Best Local Similarity 90.7%; Pred. No. 2.8e-34;
Matches 223; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

QY 658 ctctccttaactcaagccctgttcaatgtctatccaaatgttcaatgtct 717
658 ctctccttaactcaagccctgttcaatgtctatccaaatgttcaatgtct 717

Db 1 ctctccttaactcaagccctgttcaatgtctatccaaatgttcaatgtct 717
1 ctctccttaactcaagccctgttcaatgtctatccaaatgttcaatgtct 717

QY 718 acacagacatcgaggttggctgaagttaccacagttcaagatcagagaccagctt 777
718 acacagacatcgaggttggctgaagttaccacagttcaagatcagagaccagctt 777

Db 61 acacagacatcgaggttggctgaagttaccacagttcaagatcagagaccagctt 777
61 acacagacatcgaggttggctgaagttaccacagttcaagatcagagaccagctt 777

QY 778 tgggattctctctgtagtgggagatcctcactcaagaagaaggggtccagtggcacta 837
778 tgggattctctctgtagtgggagatcctcactcaagaagaaggggtccagtggcacta 837

Db 121 tgggattctctctgtagtgggagatcctcactcaagaagaaggggtccagtggcacta 837
121 tgggattctctctgtagtgggagatcctcactcaagaagaaggggtccagtggcacta 837

QY 838 ataattggaagccggtgtgcatcttgagaagcctcctaactccagttcatcatca 897
838 ataattggaagccggtgtgcatcttgagaagcctcctaactccagttcatcatca 897

Db 179 ataattggaagccggtgtgcatcttgagaagcctcctaactccagttcatcatca 897
179 ataattggaagccggtgtgcatcttgagaagcctcctaactccagttcatcatca 897

QY 898 gcaacca 903
898 gcaacca 903

Db 233 GCCCAA 238

RESULT 4
LOCUS
DEFINITION

AV527594 540 bp mRNA EST 01-SEP-2000
AV527594 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone AP240f09R 5', mRNA sequence.

ACCESSION	VERSION	AV527594	GI:8687122
KEYWORDS	EST		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 540)		
TITLE	Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.		
JOURNAL	A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized size-selected cDNA libraries		
MEDLINE	DNA Res. 7, 175-180 (2000)		
COMMENT	20363093		
FEATURES	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yakusa 1532-3, Kisarazu, Chiba 292-0812, Japan. Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/		
SOURCE	Location/Qualifiers		
	1..540		
	/organism="Arabidopsis thaliana"		
	/strain="Columbia"		
	/db_xref="taxon:3702"		
	/clone="AP240E09"		
	/clone_id="Arabidopsis thaliana aboveground organs two to six-week old"		
	/tissue_type="aboveground organs"		
	/dev_stage="two to six-week old"		
	/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"		
BASE COUNT	175 a 124 c 110 g 131 t		
ORIGIN			
Query Match	11.9%; Score 154.4; DB 31; Length 540;		
Best Local Similarity	67.0%; Pred. No. 6.9e-31;		
Matches 240; Conservative	0; Mismatches 106; Indels 12; Gaps 1		
Oy	23 aagaagatttctgtggaagccaatbaacagatctgtgacaacacatgcttaacaaca 82		
Db	151 ACAGGAGATTGTAGTGAACCCCTTGAGAGAAATAGTACAAACACATGCCCAAGCCGA 210		
Oy	83 gtcgaatctctcttcctcctcgtctactcc-----tctggaagaaggaagcaagcc 130		
Db	211 ACCGCAACCGGCTTCAACCGCAGCAGCACCGTCCGGTGGGTGAGAGGAGCAAGGCC 270		
Oy	131 ggagaagaaccagagcttgaactgtgccaaagatgcgaactctctaaacacaaatctgtta 190		
Db	271 AGAAAAGATCAAGGTGTAACTGTCCGAGATGTAACCTCAACACACAAAGTTTGTGA 330		
Oy	191 ctacacacactacaagcctcgaagcgaagccagagtaactttgtaagaagcagaagagatgtg 250		
Db	331 CTACAAACATTAATTACTTTGACGACGACCAAGATACCTTGTCCAAAGGTGTGAGAGGATTTG 390		
Oy	251 gacgcgagtggttccctccagaagaatcccggtcggtgagcgcgctcgcgaagaacaagag 310		
Db	391 GACCGAAGGCGGTTGCTTGAAGAAATTCCTTTGGCGGTGGCTCAAGAAAGAACAAAGAG 450		
Oy	311 atctcttcacaatctctctctcctcctcaaccccttcctcgtctcttccttaagaagaacct 368		
Db	451 ATCTACACTTCTTCTTCTGTGATATTAGTAAACATCACTCGATTCGATTCTACCAACCAAGCT 508		
RESULT	5		
LOCUS	AW233999	526 bp	EST
DEFINITION	sfs2h11.v1 gm-cl028 Glycine max cDNA clone GENOME SYSTEMS CLONE	17-JUL-2000	
ACCESSION	AW233999	PROTEIN	
VERSION	AW233999.1	GI:6566326	

KEYWORDS	EST.
SOURCE	soybean.
ORGANISM	Glycine max
REFERENCE	Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae: eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
AUTHORS	1 (bases 1 to 526) Shoemaker, R., Keim, P., Vodka, L., Erpelidg, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, B., Allent, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE	Public Soybean EST Project
JOURNAL	Unpublished (1999)
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com Insert Length: 1210 Std Error: 0.00 High quality sequence stop: 405.
FEATURES	location/Qualifiers 1..526 /organism="Glycine max" /db_xref="taxon:3847" /clone="GENOME SYSTEMS CLONE ID: Gm-cl028-lj26" /clone_lid="Gm-cl028" /clone_type="roots of 'Supernod' plants" /lab_host="DH10B" /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; The mRNA was isolated from roots of Glycine max 'Supernod' plants generously donated by Dr. Gary Stacey. The seedlings were inoculated with Bradyrhizobium japonicus, strain USDA110 prior to harvest. Stratagene's cDNA synthesis kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V-A-C, or G) was added to the 3' end of the primer (GAGAGAGAGAGAGAGAGAGACTGCTGAC(7118V) to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA polymerase, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript II XR predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Corryell."
BASE COUNT	144 a 146 c 102 g 134 t
ORIGIN	Query Match 11.0%; Score 142.6; DB 113; Length 526; Best Local Similarity 71.6%; Pred. No. 1.1e-27; Matches 187; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

	RESULT	6
AM597702		
LOCUS	AM597702	558 bp mRNA
DEFINITION	sJ49g95.5' Gm-cl033 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:	18-JUL-2000
	Gm-cl033-657 5'	similar to TR:082027 082027 DOF ZINC FINGER PROTEIN
	:: mRNA sequence.	

ORGANISMS

REFERENCE
AUTHORS
Shoemaker, R., Keim, P., Vodkin, L., Eprelding, J., Coryell, V., Khanna,
1 (pages 1 to 356)
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta,
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
Rosidae: eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine,
1 (pages 1 to 356)
Shoemaker, R., Keim, P., Vodkin, L., Eprelding, J., Coryell, V., Khanna,

TITLE	JOURNAL	COMMENT
Public Soybean EST project	Unpublished (1999)	Contact: Shoemaker R/Public Soybean EST project

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estlewatson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Insert Length: 1206 Std Error: 0.00
High quality sequence stop: 371.

FEATURES
source

```

/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl03-657"
/clone_lib="Gm-cl03"
/rclone_type="Desloy 5710' seedling roots"
/lab_host="Dh10a"
/notice="Vector: Bluescript II XR; Site_1: EcoRI; Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from 'Desloy 5710' seedling roots. Tissue was taken from
7-day-old seedlings that had been propagated on paper
towels with distilled water. Tissue was taken from the tip
to the first lateral root, usually about 3cm from the tip."

```

and flash-frozen in liquid nitrogen. StrataGene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of StrataGene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A, C, or G) was added to the 3' end of the primer (GAGACAGACAGACAGACAGAACTACTCTCTGAC(TT)18V) to anchor the primer at the 5' end of the poly(A) tract. After second strand synthesis, the cDNA ends were filled in with cloned pfu DNA, ligated to EcoRI adaptors and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (400u/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GIBCOBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then precipitated, redissolved, and ligated into StrataGene's pBluescript II XR predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 100% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=24 and 6, respectively). This library was constructed by Dr. Paul Kelm and Dr. Virginia Corvelli."

BASE COUNT	161 a	154 c	110 g	133 t
ORIGIN				

Query Match	Score	DB	Length
11.0%	142.6	118	558

Best Local Similarity	1.68	Pred.No.	1.1e-27
Matches	187	Conservative	0
		Mismatches	74
		Indels	0
		Gaps	0

QY 80 acagctcgaaatcctcctcctcctgctactcctctgtgaaaggaaggaacccgataaaga 135

D_b 54 AAGCCTCCCTCCTCCTCCTCCTATGTTGGAGAGAAGGGCAAGGCCTCAAAAGGA 113

Qy 140 ccagccttgaactgtccaagatgcacaccttaaacccaagttctgttactacaaca 199

Db 114 TCAAGCTTTGAACTGCCCAAGGTGCAATTCAACAAACACCAAAATTTCTGCTACTACACAA 173

QY 200 ctacagcctgacgcagccacgctactttgttaagactgcagaggtatgaccgcag 255

Db 174 CTATAGCCTCTCTCAGCCAGGTACTTTTGCAAGACATGTAGAGGATTTGCACTGAGGG 233

260 tggttccctcaggaacatccccgtctgctgctgctccgcaagaacaagatctcttc 315

Db 234 TGGTCTCTCAGAAATGTTCTGCTGGGGTGGCGCTTAGAAGACAGAGATCAACCC 2933

QY 320 caatcctcttctcttacc 340

Db 294 GCCAGCACCACATCAGCACC 314

5
4
3
2
1
E

RESULT	/
AW620835	

DEFINITION	sj47b12.y1	Gm-c1033	glycine max	cdna	clone	GENOME	SYSTEMS	CLONE	ID:
LOCUS	AM620835	448 bp	mrna	EST	18-JUL-2000				

GM-CL103-408 5' similar to TR:082027 082027 DOF ZINC FINGER PROTEIN mRNA sequence.

ACCESSION	AM620835
VERSION	AM620835.1
	GI:7327037

KEYWORDS	EST.
SOURCE	soybean.

ORGANISM Glycine max
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae: eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE	
AUTHORS	
1 (bases 1 to 448)	
Shoemaker, R., Keim, P., Vodkin, L., Erpelidng, J., Corvelli, V., Khanna	

.....

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
Populus tremula x Populus tremuloides Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Malpighiales; Salicaceae; Populus.	1 (bases 1 to 532)	Sterky,F., Regan,S., Karlsson,J., Hertzberg,M., Rohde,A., Holmberg ,A., Amali,B., Bhalero,R., Larsson,M., Villaruel,R., Van Montagu ,M., Sandberg,G., Olsson,O., Teerl,T.F., Boerjan,W., Gustafsson,P., Uhlen,M., Sundberg,B. and Lundberg,J. Gene discovery in the wood-forming tissues of poplar: Analysis of 5,692 expressed sequence tags Proc. Natl. Acad. Sci. U.S.A. 95 (22), 13330-13335 (1998) 99007314	Contact: Sterky F Department of Biotechnology Royal Institute of Technology (KTH) Teknirringen 34, S-100 44 STOCKHOLM, Sweden Tel.: +46 8 790 8287 Fax: +46 8 24 54 52 Email: fredrik@biochem.kth.se			
PCB Primers FORWARD: AAAGGGGATGTGCTGCAGGCG BACKWARD: GCTTCCGCTCGATGTCGTCG Seq primer: CGTTGTAACAGCAGCGCCAG High quality sequence stop: 532. Location/Qualifiers	1..532 /organism="Populus tremula x Populus tremuloides" /db_xref="taxon:47664" /clone_lib="Hybrid aspen plasmid library" /issue_type="Cambial region" /dev_stage="1.5 m actively growing tree" /lab_host="E.coli" /note="Vector: pBluescript SK; Site_1: SalI; Site_2: NotI; Cambial region tissues, including developing xylem, the meristematic cambial zone and the developing and mature phloem, was harvested from 1.5 m actively growing trees. cDNA was prepared and cloned into lambda gt2a. DNA was isolated and subcloned into pBluescript SK using SalI and NotI restriction enzymes."					
BASE COUNT	146 a	116 c	119 g	148 t	3 others	
ORIGIN						
Query Match	9.3%	Score 120.6;	DB: 16;	Length 532;		
Best Local Similarity	71.3%;	Pred. NO. 9.5e-22;				
Matches 159;	Conservative	0;	Mismatches 64;	Indels 0;	Gaps 0;	
OY 112	tggaaaggaagcaagacgaggagaagaagccgagcttgtaactgtccaagaatccaactcct 171					
Dd 111	TAGAAAGAAGGCCAACGCCCAAAAGGAACAAGCTTTAAATTGTCCGAGTGCAACTTGA 170					
OY 172	taaacaccagttcttgttacttaacaacctacagcctgcagcgagccagagtactttgtla 231					
Dd 171	CTAAATACAAAGTTTTGTACTACACAACTATAGTCTCTCCAGCACCAAGAAATCTTTTCCA 230					
OY 232	aagacttgagagagratitgagccgaggttgtttcctctaaggaacaatcccgcgtgtagcg 291					
Dd 231	AGACTTGTAGAAAGTACTGCAGCTGAAGGTGGGTCTTTAAGAAATGTTCTCTGTGGTGSTG 290					
OY 292	gcgtccgcaagaacaagaatatcttctccaattccactctctctc 334					
Dd 291	GTTGAGAAAGAACAGAGATCATCAAGTAATCCATCATCATCATC 333					
RESULT 14						
LOCUS	A1894846	571 bp	mRNA	EST	27-JUL-1999	
DEFINITION	ESR264289 tomato callus, TAMU Lycopersicon esculentum cDNA clone					
ACCSSION	A1894846					
VERSION	A1894846.1	GI:5600748				
KEYWORDS	EST.					

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2001, 17:00:51 ; Search time 7917.7 Seconds
(without alignments)
1457.738 Million cell updates/sec

Title: US-09-394-519-127

Perfect score: 1221
Sequence: 1 attctctccacaagaagtl.....aaaaaaaaaaaaaaaaaaaa 1221

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_est3.*

4: gb_est4.*

5: gb_est5.*

6: gb_est6.*

7: gb_est7.*

8: gb_est8.*

9: gb_est9.*

10: gb_est10.*

11: gb_est11.*

12: gb_est12.*

13: gb_est13.*

14: gb_est14.*

15: gb_est15.*

16: gb_est16.*

17: gb_est17.*

18: gb_est18.*

19: gb_est19.*

20: gb_est20.*

21: gb_est21.*

22: gb_est22.*

23: gb_est23.*

24: gb_est24.*

25: gb_est25.*

26: gb_est26.*

27: gb_est27.*

28: gb_est28.*

29: gb_est29.*

30: gb_est30.*

31: gb_est31.*

32: gb_est32.*

33: gb_est33.*

34: gb_est34.*

35: gb_est35.*

36: gb_est36.*

37: gb_est37.*

38: gb_est38.*

39: gb_est39.*

40: gb_est40.*

41: gb_est41.*

42: gb_est42.*

44: em_esthum10.*
45: em_esthum11.*
46: em_esthum12.*
47: em_esthum13.*
48: em_esthum14.*
49: em_esthum15.*
50: em_esthum16.*
51: em_esthum17.*
52: em_esthum18.*
53: em_esthum19.*
54: em_esthum20.*
55: em_esthum21.*
56: em_esthum22.*
57: em_esthum23.*
58: em_esthum24.*
59: em_esthum25.*
60: em_esthum26.*
61: em_esthum27.*
62: em_esthum28.*
63: em_esthum29.*
64: em_esthum30.*
65: em_esthum31.*
66: em_esthum32.*
67: em_esthum33.*
68: em_esthum34.*
69: em_esthum35.*
70: em_esthum36.*
71: em_esthum37.*
72: em_esthum38.*
73: em_esthum39.*
74: em_esthum40.*
75: em_esthum41.*
76: em_esthum42.*
77: em_esthum43.*
78: em_esthum44.*
79: em_esthum45.*
80: em_esthum46.*
81: em_esthum47.*
82: em_esthum48.*
83: em_esthum49.*
84: em_esthum50.*
85: em_esthum51.*
86: em_esthum52.*
87: em_esthum53.*
88: em_esthum54.*
89: em_esthum55.*
90: em_esthum56.*
91: em_esthum57.*
92: em_esthum58.*
93: em_esthum59.*
94: em_esthum60.*
95: em_esthum61.*
96: em_esthum62.*
97: em_esthum63.*
98: em_esthum64.*
99: em_esthum65.*
100: em_esthum66.*
101: em_esthum67.*
102: em_esthum68.*
103: em_esthum69.*
104: em_esthum70.*
105: em_esthum71.*
106: em_esthum72.*
107: em_esthum73.*
108: em_esthum74.*
109: em_esthum75.*
110: em_esthum76.*
111: em_esthum77.*
112: em_esthum78.*
113: em_esthum79.*
114: em_esthum80.*
115: em_esthum81.*
116: em_esthum82.*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Query Match	Best Local Similarity	Matches	Score	DB	Length
39.8%	97.5%	503	485.4	104	526
Conservative	0	Mismatches	12	Indels	1
Gaps	1				
1	attctctccacaagaagtcctaacttcgaagtgtaacaaacaccattctcatctta	60			
1	ATTCTCTCCACAAAGAAGTCCTACTTCGAGTTGAACAAACACCATTTCTCATCTCA	60			
61	tctcagaagaagacaacacatttcgtgtctctctcttctctatctccataaggaatat	120			
61	TCTCGAAGAAGAACAAACCATTTCTGTTCTTCTTCTCATTCATCAAGAAATTA	120			
121	ttccggaacagctgagcttcgttctgtaagaagaataaacaatgatagtggcaagaag	180			
121	TTCTCGAAGACTGTGAGTTCTTGTGCAAGGAATTAATAAACATGATGGCAAGAG	180			
181	atctaggttcagccctaagcttaagggttttcacaaatcacacatctcttcagatga	240			
181	ATCTAGGTTTGAGCCTTAGGCTTAGGGTTTTCACAAATCACAAATCTCTTCAGATGA	240			
241	tgaatccctaactctcatatatacaacaacatctccagagatcccatlgagaacaa	300			
241	TGAATCCCTACTCTCATCATCAACAAATCTCCAGAGACTCCCATGGAACCAACATTCG	300			
301	atccctaactcagatcttcgcaagaataagcgtgaaagtttccatcaacagttact	360			
301	ATCCCTAATCAGATCTTTCGCAAGATAGCGTGAACAGTTTTCATCAACAGGTTAATCTCG	360			

LOCUS	DEFINITION	SEQUENCE
LOCUS	743527	538 bp mRNA EST 07-JAN-1998
DEFINITION	6790 Lambda-PRL2 Arabidopsis thaliana cDNA clone 120N14T7, mRNA	
SEQUENCE		

ACCESSION	T43527
VERSION	T43527.1
KEYWORDS	GI-2758376
SOURCE	EST.
ORGANISM	thale cress. Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi. N (bases 1 to 538) Newman,T., deBrujin,F.J., Green,P., Keestra,K., Kende,H., McIntosh ,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel ,E. and Somerville,C.
AUTHORS	
TITLE	Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994)
JOURNAL	On Jan 7, 1998 this sequence version replaced gi:947926.
MEDLINE	Contact: Thomas Newman
COMMENT	MSU-DOE Plant Research Laboratory Michigan State University MSU-DOE-PTL, Michigan State University, Plant Biology Bldg., E. Lansing, MI Tel.: 517-353-0854 Fax: 517-353-9168 Email: 22313jcn@ibm.cl.msu.edu Seq primer: T7.
FEATURES	Location/Qualifiers
source	1..538 /organism="Arabidopsis thaliana" /strain="var columbia" /db.xref="taxon:3702" /clone="12ON14T7" /clone_1lb="Lambda-PRL2" /note="Vector: lambda Z1p-lox; Site_1: Sal; Site_2: Not; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark-rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and silicles). The vector is BR's lambda Z1p-lox. The cDNA inserts were directionally cloned with Sal-I Not arms using oligo dt primed cDNA."
BASE COUNT	143 a 114 c 129 g 131 t 21 others
ORIGIN	
Query Match	31.2%; Score 380.4; DB 189; Length 538;
Best Local Similarity	93.3%; Pred. No. 1.2e-84;
Matches 434; Conservative 0; Mismatches 27; Indels 4; Gaps 4;	
1 atttcttcaccaagaagtcctaacttcgaattgaaacaacaccatctcatccta 60	
2 db 1 atttcttcaccaagaagtcctaacttcgaattgaaacaacaccatctcatccta 60	
3 61 tctcgaaagaacaacaaccttcgtgttccttccttcctatctcataaggaaataa 120	
4 db 61 tctcgaaagaacaacaaccttcgtgttccttccttcctatctcataaggaaataa 120	
5 61 tctcgaaagaacaacaaccttcgtgttccttccttcctatctcataaggaaataa 120	
6 db 61 tctcgaaagaacaacaaccttcgtgttccttccttcctatctcataaggaaataa 120	
7 121 ttctcgaacagtgttgattcttgtaaaggaaataaaaactgatgtggcaagaag 180	
8 db 121 ttctcgaacagtgttgattcttgtaaaggaaataaaaactgatgtggcaagaag 180	
9 121 ttctcgaacagtgttgattcttgtaaaggaaataaaaactgatgtggcaagaag 180	
10 db 121 ttctcgaacagtgttgattcttgtaaaggaaataaaaactgatgtggcaagaag 180	
11 161 atcttaggtttgagcctaagcttaggttttccaacaatccaatcccttcgaatgaac 240	
12 db 161 atcttaggtttgagcctaagcttaggttttccaacaatccaatcccttcgaatgaac 240	
13 161 atcttaggtttgagcctaagcttaggttttccaacaatccaatcccttcgaatgaac 240	
14 db 161 atcttaggtttgagcctaagcttaggttttccaacaatccaatcccttcgaatgaac 240	
15 241 tgaatcccnaactctcatatcaacaacatctccagaagatcccatggaacccaacatcg 300	
16 db 241 tgaatcccnaactctcatatcaacaacatctccagaagatcccatggaacccaacatcg 300	
17 241 tgaatcccnaactctcatatcaacaacatctccagaagatcccatggaacccaacatcg 300	
18 db 241 tgaatcccnaactctcatatcaacaacatctccagaagatcccatggaacccaacatcg 300	
19 301 atccctcacatgcagatcttcgcaagaatagac-gtgaacagtttccatcaaac-gttaactg 358	
20 db 301 atccctcacatgcagatcttcgcaagaatagac-gtgaacagtttccatcaaac-gttaactg 358	

[illegible]

Db 141 AGACCCGGAGACTCCAG-AAAGGCGTATGAGCTTCCAGCTCTAAGCTGTCTCCACAT 199
Oy 832 tctatgctcagatbactcaccacactaacatcatcattgctctcgtgtagcgtag 891
Db 200 TCTACGGTCAGATGACTCTCCACCACTACACTCATCATGTGTCTTCGTCGACGAGTGG 259
Oy 892 ctgtgctcatcatcagacatcacacacaacacagcgcttgcattacaccgttga 951
Db 260 GTGGCCATCATCATCGAAGCATCACACCAATCATACAGCGCCGTTTCATCATCCGTGG 319
Oy 952 ttgctgtgctgctcaggttggtcattggtcgtgaatttgaagccttgcgtccagatcgt 1011
Db 320 TTGCTGTGCTGCTGTCAGGTGGCTCATGGCGCAATTTTGAAGCCTTGCTCAGATCGT 379
Oy 1012 aat-----ttttatgtgtggtggaaggggtgttttgggtttttcatatcgttatag 1065
Db 380 GATTTTATTATTATTAGTGGTGGGAAAAGGGGTGTTTGGTATTGTTTCGTTATGTAATAG 439
Oy 1066 tctatcgtctgtggtggtcattgtaatttggatgtattggtccttccatgaaatagcat 1125
Db 440 TCTATCTGTGAGGGGTCAATGTAATTTGGATGATGTGGCTTCATCATGACTAGCTAT 499
Oy 1126 gtagatgcaaccttaaatatttcaagtgcacaaatattac 1171
Db 500 GTATGATGCAACCTTAAAAAGATTAAATTGACNAAAATTGATTAC 545

RESULT 5
T42030 LOCUS T42030 475 bp mRNA EST 07-JAN-1998
DEFINITION 5293 Lambda-PRL2 Arabidopsis thaliana cDNA clone 110C167, mRNA sequence.
T42030
ACCESSION T42030
VERSION T42030.1 GI:932917
KEYWORDS EST.
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eunotids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 475)
Newman,T., deBuijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlgege,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994)
95148729
On Nov 29, 1993 this sequence version replaced gi:634618.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcm@bm.cl.msu.edu
Seq primer: T7.
Location/Qualifiers
1. 475
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="110C167"
/clone_1b="Lambda-PRL2"
/note="Vector: lambda zip-lox; site_1: Sal; site_2: Not; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark-16 hours; 4) same plants as 3 but aerial tissue (stems, flowers and

JOURNAL
MEDLINE
COMMENT

Large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)

Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22333cne@lm.cl.msu.edu

Seq primer: 17 dye primer.
Location/Qualifiers

FEATURES
source

1. 390
/organism="Arabidopsis thaliana"
/strain="var. columbiana"
/db_xref="taxon:3702"
/clone="200B97"
/clone_1lb="Lambda-PRU2"
/note="Vector: lambda Z1P-10x; Site_1: Sal; Site_2: Not;
Lambda PRU2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRU's lambda Z1P-10x. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dT primed cDNA."

BASE COUNT 94 a 79 c 95 g 107 t 15 others
ORIGIN

Query Match 23.5%; Score 287; DB 158; Length 390;
Best Local Similarity 86.7%; Pred. No. 2.5e-61;
Matches 333; Conservative 0; Mismatches 43; Indels 8; Gaps 2;

QY 705 accaagttaaagcaacgagtagatctgcgaacttgaaacgctgctagagaagct 764
|||||
Db 1 ACCAAGTTAAGCAACGAGGAGTATGCAATTAAGTAAAGGCGTACGAGCAACGTA 60
QY 765 acggaagagaaacggaacttcagaaagagctatgagagcttcgaactcctaagctgtct 824
|||||
Db 61 ACAGGAAGANAACCGGAGCTTCAGAAAGAGCGTATGAGCTTCGAACCTCAAGCTGCT 120
QY 825 ccaacattctacgctgtagatgtagctacccaactacactatcgtgtccttgctgctgag 884
|||||
Db 121 CCACAAATTNTGCGTACATGACTCCACCAACTACACTCATCTATGATGCTGCGAG 180
QY 885 cgtgtagctgtgcatcattcgaacacatcaccacactacactatcgtgtccttgctgctgag 944
|||||
Db 181 CMTNTGGTGGCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 240
QY 945 ccgtgtagctgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1003
|||||
Db 241 CCCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 1004 acgactgtaattttt-----agtgctgagggagagagtgctgtgtgtttttcattatc 1056
|||||
Db 301 ACAGNTGCAATTTTNTTNTTAAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
QY 1057 gttatatagctatcgtgtgaggg 1080
|||||
Db 361 GNAANATGCTTNTCTTGTGGG 384

RESULT 10
LOCUS BG441022 860 bp mRNA EST 15-MAR-2001
DEFINITION GA_Ea0011P23f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_Ea0011P23f, mRNA sequence.
ACCESSION BG441022
VERSION BG441022.1 GI:13350674

KEYWORDS
SOURCE
ORGANISM

EST.
Gossypium arboreum.
Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eudicots II; Malvales; Malvaceae; Gossypium.

REFERENCE
AUTHORS
TITLE

Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, D., Henry
D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATGACTCACTATGAGG
High quality sequence stop: 746.

FEATURES
source

1. 860
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0011P23f"
/clone_1lb="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 257 a 190 c 222 g 186 t 5 others
ORIGIN

Query Match 19.9%; Score 243.4; DB 153; Length 860;
Best Local Similarity 67.3%; Pred. No. 2.4e-50;
Matches 406; Conservative 0; Mismatches 161; Indels 36; Gaps 3;

QY 299 cgatcctacatcagatcttcgaagatagacgtgaaagcttccatcaacggttaactg 358
|||||
Db 7 CGAGACTATGATCTTTTAAAGAGATTCAGACGAGATTCGATTCGATTCGATTCGATTC 66
QY 359 cgaaggaagcaacgaggttctgcaccacaacagtagatcctaagcaacttagcggagaa 418
|||||
Db 67 CGAAGAGAGAGCTGGAGTTTCACTTCGAAACAGCAGATATGAG--TGTGAGTGGGAA 123
QY 419 gagaagctgagaagaagaatctcggaaacggcgcttgctccggcagcatcagcaga 478
|||||
Db 124 AAGGAGCGAAGGAGGAGCGCATATGGA-----GATCA 156
QY 479 gatactccggatcgaaggtaactacgctggaactatagaagaagaagcggggcga 538
|||||
Db 157 GCTTGATATTTAAAGAGCTGTTCTGCTGTCATCATGATGAGGAGAT-----GCTGA 210
QY 539 aacgtgagaagaagaactaggtatcacaagaatagctgtgtcttctcgaagaagactt 598
|||||
Db 211 CGCTTGAGAGAAAAAATTAAAGCTTTTCGAAAGATCAGTCTGCTATTTCTTGAAGAAAGCTT 270
QY 599 caaagaacacacactcctaactcccaacagaagctagcttggcttaagaagctgaactt 658
|||||
Db 271 TAAAGAACAAACACTCTGAAACCAAAAGAGAGGCTTTGGCTTAACAGCTGGGATT 330
QY 659 gacgcaagacaagtggaagtgtgtccaaaacagaagagctagaaccaagttaagca 718
|||||
Db 331 GCGAGCCAGACAGATTGAAGTTGTTCCAAAACAGAGGCGAGGAGCCAGCTGAGCA 390
QY 719 aacgagagtagattggaagacttgaacggtgagcgaagaagcgaagaagcgaagcga 778
|||||
Db 391 AACTGAGGTTGACTGTGAGTTCTTAAGAGAGATGCTGTGAGAAATCTGACGAGGAGAAACAG 450

QY 779 gagacttcagaagaagcctatgagcttcgaactcgaagctgtctcccaattctacgg 838.
 DB 451 GCGGTTCCAGAAAGAACTTAAAGAGCTGAGACCTGAAACTTCTCCTCAGTTCTACAT 510
 QY 839 tcagatgactccaccactacatcatgctgtccttcgtgagagtgtagtgc 898
 DB 511 GCAAAATGACCCACCTACGACCTCACCAGTGTGCCCATGTGTGAGAGAGTAGTGATC 570
 QY 899 atc 901
 DB 571 ATC 573

RESULT 11
 AM727255 standard; RNA; EST; 868 BP.
 ID AM727255
 AC AM727255;
 SV AN727255.1
 DT 24-APR-2000 (Rel. 63, Created)
 DT 18-NOV-2000 (Rel. 65, Last updated, Version 2)
 DE GA_Ea001f23 Gossypium arboreum 7-10 dpa fiber library Gossypium arboreum
 DE cDNA clone GA_Ea001f23, mRNA sequence.
 XX EST.
 KM Gossypium arboreum
 OS Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Malvales; Malvaceae; Gossypium.
 OC [1]
 RN 1-868
 RP Wing R.A., Frisch D., Yu Y., Main D., Rambo T., Simmons J., Henry D.,
 RA Wood T.C., Leslie A., Wilkins T.A.;
 RT "An integrated analysis of the genetics, development, and evolution of the
 RT cotton fiber";
 RL Unpublished.
 CC Contact: Wing RA
 CC Clemson University Genomics Institute
 CC Clemson University
 CC 100 Jordan Hall, Clemson, SC 29634, USA
 CC Tel: 864 656-7288
 CC Fax: 864 656-4293
 CC Email: rwing@clemson.edu
 CC High quality sequence stop: 868.
 CC XX
 FH Key Location/Qualifiers
 FH source 1. 868
 FT /db_xref="taxon:29729"
 FT /db_xref="ESTLIB:2480"
 FT /note="Vector: PBK-CMV; Site_1: EcoRI; Site_2: XhoI"
 FT /organism="Gossypium arboreum"
 FT /strain="AKA"
 FT /cultivar="8400"
 FT /clone="GA_Ea001f23"
 FT /tissue="Gossypium arboreum 7-10 dpa fiber library"
 FT /tissue_type="Fibers isolated from bolls harvested 7-10
 FT dpa"
 FT /lab_host="E. coli"
 FT XX
 SQ Sequence 868 BP; 259 A; 192 C; 226 G; 186 T; 5 other;

QY 299 cgatccacatcagatcttcgcaagatagacgtgaacagtttccatcaacggttaactg 358
 DB 15 CGAAGCTAGATCATTTCTTAGAGAAATGACGTGAACAGATGTGCATCTACGCGGATG 74
 QY 359 cgaaggaagacaagaggttccgttcacaaacagatagatctcaagccattagcgagaa 418
 DB 75 CGAAGAAAGAAAGCTGAGATTTTCATCTCCGAACAGACATATTCAG---TGTGATGGGAA 131
 QY 419 gagaagtgaagagaagaaatctccggaacgagctgtgcgcgagacatcaagca 478
 DB 132 AAGGAGCGAAAGGAAAGCAAGCATATGCA-----GATGA 164
 QY 479 gatactccgagatcggaggtactcaagtggaacctcagatgaagaagaacgaggcga 538
 DB 165 GCTTGATATTGTAAGAGCTTGTCTCCTGCGCATCAGATGATGAGAAAT-----GGTGA 218
 QY 539 aagtcgagagaagaagctcggtatcaaaagatcagctgtcttctcgaagaagactt 598
 DB 219 CGCTTCGAGAAAGAAAGCTTAAAGCTTTCGAAGATCAGTGTCTATTCTTGAGAAAGCTT 278
 QY 599 caaagaacacacacatctcaatcccaacagaagaagctagcttggtctaagaagctgaact 658
 DB 279 TAAAGAACACACACCTCTGAACCCAAACAAAGATGGCTTGGCTAAGCAGCTGGGATT 338
 QY 659 gaagcaagacaagtggaagtgtgtccaaacagaagaagctagaacccaagttaagca 718
 DB 339 GCGACCCAGCAAGTGAAGTTGTGTTCCAAACAGAAAGGCGGAGACCAAGCTGAAGCA 398
 QY 719 aacgaggtagatgtagaatacttgaacggtgcgtgtagaagaagctagaagaacg 778
 DB 399 AACTGAGGTTGACTGTGATGTTCTTAAGACATGCTGTGACATATGACGAGGAAAAACAG 458
 QY 779 gagacttcagaagaagcctatgagcttcgaactcgaactcgtctcccaattctacgg 838
 DB 459 GCGGTTGAGAGAGAGAGTTCAAGAGCTGAGAGCAGCTGAACCTTCTCCTCAGTTCTACAT 518
 QY 839 tcagatgactccaccactacatcatatgcttcttcgtgagagtgtagtgc 898
 DB 519 GCAAAATGACCCACCTACGACCTCACCAGTGTGCCCATGTGTGAGAGAGTAGTGATC 578
 QY 899 atc 901
 DB 579 ATC 581

RESULT 12
 AM277753
 LOCUS 607 bp mRNA EST 18-APR-2000
 DEFINITION s186508.y1 Gm-c1019 glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1019-2992 5' similar to SW:ATHA_ARATH_P92953 HOMEBOX-LEUCINE ZIPPER PROTEIN ATHB-4 ;, mRNA sequence.
 ACCESSION AM277753
 VERSION AM277753.1 GI:6666294
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
 REFERENCE 1 (bases 1 to 607)
 AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Rhana
 A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
 Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers
 Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
 R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cadenas,M., McCann
 R., Waterston,R. and Wilson,R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished (1999)
 COMMENT Public Soybean EST Project
 Public Soybean EST Project
 Contact: Shoemaker R/Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2001, 10:45:17 ; Search time 7917.7 Seconds
(without alignments)
1875.599 Million cell updates/sec

Title: US-09-394-519-19

Perfect score: 1571

Sequence: 1 aggaacagtgaaaggttcg.....gaagtttagttatgatttg 1571

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	483.8	30.8	466	256	B29089 T26B14TR TA
2	417	26.5	417	256	B77441 T26B14TR TA
3	394.4	25.1	404	256	B27142 T29H6TR TAM
4	364	23.2	438	256	N65486 20526 Lambd
5	358	22.8	358	219	CNS000W4
6	354.6	22.6	758	256	B20632 Arabidops
7	313	19.9	436	158	H76458 18163 Lambd
8	301	19.2	304	157	F14268 ATTS5190 G1
9	301	19.2	487	31	AV533904 AV533904
10	168	10.7	231	219	CNS000W2C
11	166	10.5	168	219	CNS000WNP
12	165.4	10.5	213	191	Z34128 Arabidops
13	145.2	9.2	763	256	B20629 Arabidops
14	138	8.8	410	136	BE524424 M4969STM
15	119.6	7.6	492	146	BF279235 GA_Eb003
16	85	5.4	486	155	BG596420 EST495098
17	85	5.4	492	110	AM029972 EST273227
18	85	5.4	494	110	AM033481 EST277052
19	85	5.4	511	113	AM217152 EST295866
20	85	5.4	513	118	AM623005 EST320950
21	85	5.4	518	110	AM034705 EST278507
22	85	5.4	605	118	AM625619 EST319526
23	85	5.4	717	155	BG591978 EST499820
24	81.8	5.2	462	123	AM979604 EST341203
25	80.8	5.1	760	155	BG592817 EST491595
26	79.2	5.0	510	143	BF050737 EST3435895
27	79.2	5.0	515	108	AI896872 EST266315
28	79.2	5.0	545	113	AM184459 EST320445
29	79.2	5.0	564	24	AI778378 EST292527
30	79.2	5.0	564	103	AI895484 EST264627
31	79.2	5.0	598	122	AM930485 EST340858
32	79.2	5.0	601	103	AI895984 EST265427
33	79.2	5.0	742	174	BG123412 EST469058
34	77.6	4.9	561	110	AM035085 EST279913
35	77.2	4.9	433	113	AM256333 EST304540
36	77.2	4.9	478	165	BE247855 NF022B01D
37	77.2	4.9	558	150	BF520143 EST3457612
38	77.2	4.9	559	120	AM775766 EST45831
39	77.2	4.9	612	155	BG583349 EST485100
40	77.2	4.9	666	152	BG350434 091E081Ma
41	77.2	4.9	681	155	BG598073 EST496751
42	77.2	4.9	688	151	BF632128 NF025D10D
43	77.2	4.9	541	110	AM032669 EST276228
44	76.8	4.9	421	167	BE458391 EST413663
45	76.8	4.9	503	111	AM096359 EST289539

ALIGNMENTS

RESULT	1	496 bp	DNA	GSS	13-OCT-1997
LOCUS	B29089	T26B14TR	TAMU Arabidopsis thaliana genomic clone T26B14	DNA	
DEFINITION	B29089	sequence.			
ACCESSION	B29089	GI:2515055			
VERSION	B29089.1				
KEYWORDS	GSS.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
AUTHORS	Rounsley, S.D., Kelley, J.M., Field, C.E., Craven, M.B., Adams, M.D. and Venter, J.C.				
TITLE	Use of a BAC End Sequence Database To Identify Minimal Overlaps for				

JOURNAL COMMENT

Arabidopsis Genomic Sequencing
 Unpublished (1997)
 Other GSSs: T26B14TRC
 Contact: Steve Rounsley
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: rounsley@tigr.org
 Seq primer: M13 Reverse
 Class: BAC ends
 High quality sequence stop: 496.
 Location/Qualifiers
 1. 496
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="T26B14"
 /clone_lib="TAMU"
 /sex="hermaphrodite"
 /note="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII
 ; Produced by Rod Wing"

FEATURES

source

BASE COUNT

113 a 118 c 100 g 165 t

Query Match

Best Local Similarity 30.8%; Score 483.8; DB 256; Length 496;
 Matches 485; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	903	ttgatgtctatgttgcgtgaagatgtgtcactactaaacgtgatgttcttcatctgaga	962
DB	1	TTGATGCTATGTGTCGTGAATGTCACCTACTAATACGTGATGTTCTTCATCTGAGA	60
QY	963	ctaaccaatgtctcgttcttccactgttcttctgttggacgaagatcttctgcat	1022
DB	61	CTAACCAATGCTCTCGTCTTCCACTGTGTTCTGTTGACCAAGATGACACTTCTGCAT	120
QY	1023	cagcttcaacttgcgttcaacaacccatgtacgtctgacccgttgcctcaactgtccaa	1082
DB	121	CAGCTCATTGTTGTGTCAACACCTGATGACGTCGACCGCTTGCCTCAACTGCTCAA	180
QY	1083	ctccaatgttccgtcgttgcgttgcgaagaacgaagatgttgcattcgaattactaalc	1142
DB	181	CTCCAAATGTTCTCGTGTGGAAGAAGAAAGTTCGATTTGCACTTACTAATC	240
QY	1143	taacgatccctgatttggtttcttgcgaagagacgaagacactgacttgcattgtt	1202
DB	241	TACAGATCCCTGATTTGGTTTGGCGAGGAGCAACAGACCTTACATTCATTGTT	300
QY	1203	tcctcgcgagatcgaattgatattcgtcgttgcgttgcattgacattcaagattcgaag	1262
DB	301	TCCTCGCGAGATGATCAGTTTATGATTTTCGCTTGTGATGACATTCAGATTCGAG	360
QY	1263	ataacggtccagtcgttaccagattcgaattcgttgcgaatgttgaagatctcagctag	1322
DB	361	ATAACGGTCCAGTGCCTTACAGATTTGCACTTGGCGAGTTCAGATTCAGCTAG	420
QY	1323	ctgactctagtttgcgttcttcttgcataactgtctctataaactcttgcacataa	1382
DB	421	CTGACTCTAGTTTGGTTCCTTGTATCAACTTGTCTTATCAACATCTTGCACATTAA	480
QY	1383	aaattt 1389	
DB	481	AAATTTT 487	

RESULT	2	417 bp	DNA	GSS	16-JAN-1998
LOCUS	B77441	T26B16TR	TAMU Arabidopsis thaliana genomic clone T26B16	DNA	
DEFINITION	B77441	sequence.			

ACCESSION B77441
VERSION B77441.1 GI:2774080
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 417)
AUTHORS Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K., Golden,K.,
Berry,K., Granger,D., Sun,E., Wible,C., Adams,M.D. and Venter,
J.C.
TITLE A BAC End Sequence Database for Identifying Minimal Overlaps in
Arabidopsis Genomic Sequencing. Update 3
JOURNAL Unpublished (1997)
COMMENT Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 417.
Location/Qualifiers
1. 417
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="T26E16"
/clone_1ib="TAMU"
/sex="hermaphrodite"
/note="Vector: BeloBACII; site_1: HindIII; site_2: HindIII
; Produced by Rod Wing"
BASE COUNT 96 a 96 c 90 g 135 t
ORIGIN

Query Match 26.5%; Score 417; DB 256; Length 417;
Best Local Similarity 100.0%; Pred. No. 4.5e-95;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 900 agtttgatgctatgttctgtggaatgtgtccactactaaacgtgattcttctcatctg 959
Db 1 AGTTGATGCTANTGTTGCTGGAAATGTGCACCTACTAAAGCTGATGTTCTTCAATCG 60

QY 960 agactagccaagctctgttcttctacccgtgtgttccctgttgaggaagatgacattctg 1019
Db 61 AGACTAGCCAAATGCTCTGCTCTTTCACCTGTTCTCTTGTAGCAAGATGACACTTCTG 120

QY 1020 catcagctctcaactgtgttcaacaaccctgatacgtctcagaccgttctcacaatgctc 1079
Db 121 CATCAGCTCTCACTTGTGTCAACACCCGTGATGACGTCTCGACCGTTGCTCCAACTGCTC 180

QY 1080 caactcacaatgtcttcgtgtgtggaacaagaacgltgttctcgaattctcgaattacta 1139
Db 181 CAACCTCAATATGTTCTGTGCTGGGAACAAGAAAGTTTGCATTTGCACTTTACTA 240

QY 1140 atctacagatccctgatttctgttcttgaggaagagacaagaactagactcgatt 1199
Db 241 ATCTACAGATCCCTGATTTTGGTTCTTGTGCAAGAGACACTTGTGATTTGATTT 300

QY 1200 gtctctcgcgagtgatcagttgattgatttgcgtgtctgtgtgtagcattcaagattcg 1259
Db 301 GTTCTCTCGGGAGATCACTTGTGATTTTGGCTTGTGATTCAGATTCAGAGATTCG 360

QY 1260 aagataacggtccaaagtgcgttacacagatttcgaacttgcgagtggtgaaagacttc 1316
Db 361 AAGATAACGGTCCAAAGTGCCTTACGATTTGCACTTTGGGAGATGTTGAAGATCTTC 417

RESULT 3

B27142 404 bp DNA GSS 13-OCT-1997
LOCUS T29H6TR TAMU Arabidopsis thaliana genomic clone T29H6. DNA
DEFINITION sequence.
ACCESSION B27142
VERSION B27142.1 GI:2513108
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 404)
AUTHORS Rounsley,S.D., Kelley,J.M., Field,C.E., Craven,M.B., Adams,M.D. and
Venter,J.C.
TITLE Use of a BAC End Sequence Database To Identify Minimal Overlaps for
Arabidopsis Genomic Sequencing
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: T29H6TFB
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 404.
Location/Qualifiers
1. 404
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="T29H6"
/clone_1ib="TAMU"
/sex="hermaphrodite"
/note="Vector: BeloBACII; site_1: HindIII; site_2: HindIII
; Produced by Rod Wing"
BASE COUNT 91 a 98 c 83 g 132 t
ORIGIN

Query Match 25.1%; Score 394.4; DB 256; Length 404;
Best Local Similarity 98.5%; Pred. No. 2.4e-89;
Matches 398; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 895 gcttgagtttgatgcatgtgtgtggaatgtgtccactactaaacgtgattcttctc 954
Db 1 GCTTGAGTTTGATGATGCTATGTTGCTGGAATGTGTCACCTACTAAACGTGATGTTCTTC 60

QY 955 atctgagactaagccaatgctcgttcttcaacctgtgttctcgtgtgagacaatgacac 1014
Db 61 ATCTGAGACTAGCCAAATGCTCTGCTTCTTTCACCTGTTGTTCCGTGAGACATGACAC 120

QY 1015 ttctgatacagctctcaactgtgttcaacaaccctgatacgtctcgaacgttctcacaac 1074
Db 121 TTCTGCATCAGCTCTCAGTTGTGTCAACAACCTGATGAGTTCGACCGTTGCTCCAAC 180

QY 1075 tgcctcaactccaatgtcttcgtgtgtggaacaagaagatgttctgattctgact 1134
Db 181 TGCTCCAACTCCAAATGTTCTGCTGTGGAACAACGAAACGTTGTTGATTTGCACTT 240

QY 1135 tactaatctacagatccctgatttggttcttgaggaaggagacaagaactagactt 1194
Db 241 TACTAATCTACAGATCCCTGATTTTGGTTCTTGTGCAAGAGACTTCAGATTCAGATT 300

QY 1195 cgattgttctcgcgagtgatcagtttgatgattcgcgcttgctgtagatcaatcaag 1254
Db 301 CGATTGTTCTCTCGCGAGATGATGATTTGATTTCCGTTGCTTGTGATGACATTCGAAG 360

QY 1255 atcgaagaataacggtccaaagtgcgttacacagatttcgaacttgc 1298

Db 361 ATTGCAAGATACGGTCCAGTGGCTTACAGATTCCGACTTTG 404

RESULT 4
LOCUS N65486 438 bp mRNA EST 05-JAN-1998
DEFINITION 20526 lambda-PRL2 Arabidopsis thaliana cDNA clone Z30D2177, mRNA sequence.
ACCESSION N65486
VERSION N65486
KEYWORDS EST.
SOURCE GI:1217112
ORGANISM thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 438)
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E., and Somerville,C.
Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994)
95148729
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9188
Email: 22313c@edim.c1.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers
1. 438
/organism="Arabidopsis thaliana"
/strain="Var Columbia"
/db_xref="taxon:3702"
/clone="Z30D2177"
/note="Vector: lambda-PRL2"
/note="Vector: lambda Zip-Lox: Site_1: Sal; Site_2: Not; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRU's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."

BASE COUNT 99 a 90 c 98 g 134 t 17 others
ORIGIN

Query Match 23.2%; Score 364; DB 159; Length 438;
Best Local Similarity 90.1%; Pred. No. 1.2e-81;
Matches 391; Conservative 0; Mismatches 41; Indels 2; Gaps 1;

QY 790 ggggaatggcgtcgtgagatagatcctataagaanaactggactgtgtgggtac 849
|||||
Db 1 GGGGAATGGGCTGCTGAGATTAGATCCCTATTAAAGAACTAGACTGTGGGTAC 60
|||||

QY 850 ttgtgactcttgaagaagctgtaagctatgatgcttaagaagcttgagttgatgc 909
|||||
Db 61 TTTTAATACCTTTGAGAGAGCTGCTAAAGCTTAATNATGCTAAGAGCTTGATTNATGC 120
|||||

QY 910 tatgttgcgtgaagaatgtgtccactactaaacgtgatgttcttcaatctgagactagcca 969
|||||
Db 121 TATGTGCTGCGAATATGTCACACNACATAAGATNATGTTCTNCAATGTGAGACTAGCA 180
|||||

QY 970 atgctctgctctcactcactgtgttctctgttgagcaagaatgacactctgacatgctct 1029
|||||
Db 181 ATGCTCTGCTCTTACACTGTTGTCCTGTTGAGCAGATGACACTTCTGCATGACTCT 240
|||||

QY 1030 cacttgtgtcaacaaccctgtatgacgtctcagaccgttgcctccaactgtcctcaactc 1089
|||||
Db 241 CACTTGTCTCAACAACCCGTATGAGCTGTGACCGCTTCTCCCAACNGCTCCACACTNCAA 300
|||||

QY 1090 tgtctcgtcgtgtggaacaagaagactgttgcatttcgac--ttactaatctacag 1147
|||||
Db 301 TGTTCCTGCTGCTGGAACAAGAAACGTTGCTGCTTGTGCTTACTAATNCTAG 360
|||||

QY 1148 atccctgatttggcttctcagagagcaagaacactagacttcgatttctctc 1207
|||||
Db 361 NTCCCTGTTTGGGTTTGTGCGAGAGGGGCCAACAAGNCTTGGCCTTCGTTGTTTCT 420
|||||

QY 1208 gcgagatgactcagt 1221
|||
Db 421 NGCGGTNANCACTT 434
|||||

RESULT 5
LOCUS CNS00VW4/c 358 bp DNA GSS 28-JUN-1999
DEFINITION CNS00VW4 Arabidopsis thaliana genome survey sequence T7 end of BAC T10J21 of TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.
ACCESSION AL092818
VERSION AL092818.1 GI:5293972
KEYWORDS GSS.
SOURCE thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 358)
Salanoubat,M., Choisy,N., Artiguenave,F., Brothier,P., Wincker,P., Samson,D., Saurin,M., Weissenbach,J. and Queller,F.
Unpublished
2 (bases 1 to 358)
REFERENCE
JOURNAL Direct Submission
AUTHORS Genoscope.
TITLE Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191, 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Location/Qualifiers
1. 358
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_1lb="TAMU"
/clone="T10J21"
/note="end : T7"

BASE COUNT 106 a 92 c 58 g 102 t
ORIGIN

Query Match 22.8%; Score 358; DB 219; Length 358;
Best Local Similarity 100.0%; Pred. No. 3.9e-80;
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 535 gaggaaagtcctatctatgtgaatgacatcttactgactgaagatccctcctagtatga 594
|||||
Db 358 GAGGAAGTCCGTAATTATTTGATGATGATCTTATGCTACTGATATCTCTAGTAGGA 299
|||||

QY 595 ggaagagcttaagtttcttaagccaagaagaatgaagaatgcttgcttgatgaactt 654
|||||
Db 298 GGAAGAGCTTAAGGTTCTTAAGCCAAAGAAATGAAGATGATCGTCTGTGAGATTAACTT 239
|||||

QY 655 tccctctatggaatcttctgacagcctctgagagttctctcactagaagaactactaaac 714
|||||
Db 238 TCCCTTATGGAATTTTGTGAACAGCCTTGTGAGAGTTCTTCTCAGACATGACTAAAC 179
|||||

QY 715 tgatgcaagaatagctgtgtcagcttctcctcgtcgttctcctagaagaagcctgttggt 774
|||||
Db 178 TGATGCAAGATATACCTGTGTGAGCTTCTCCTGCTTCTAGGAAGAGCCTGTTGGTGT 119
|||||

QY 775 taaggcaagaagaatgggggaatggctgctgagattagagatcctattagaagaactag 834
 |||
 Db 118 TAGGCAAGAAGAAATGGGGGAATGGCTGCTGAGATTAGATCTATTAGAAAACCTAG 59
 |||

QY 835 gacttggttggtgactttgatatactcttgaagaagctgtcaagctatgatctaaag 892
 |||
 Db 58 GACTTGTTGGGTACTTTGATACCTCTTGAAAGAGCTGCTAAAGCTTATGATGCTAAG 1
 |||

RESULT 6
 B20632 758 bp DNA GSS 16-SEP-1997
 LOCUS T26E16-Sp6 TAMU Arabidopsis thaliana genomic clone T26E16, DNA
 DEFINITION sequence.
 ACCESSION B20632
 VERSION B20632.1 GI:2395686
 KEYWORDS GSS.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 758)
 REFERENCE Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
 Ecker,J.
 BAC End Sequences at ATGC
 Unpublished (1997)
 TITLE Other-GSSs: T26E16-T7
 JOURNAL Contact: Ecker J.
 COMMENT Arabidopsis Thaliana Genome Center
 University of Pennsylvania
 Dept. of Biology, University of Pennsylvania, Philadelphia, PA
 19104
 Tel: 215-898-9384
 Email: jecker@atgenom.bio.upenn.edu
 Seg primer: Sp6
 Class: BAC ends
 High quality sequence start: 192
 High quality sequence stop: 293.

FEATURES
 source
 1..758
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone_1ib="T26E16"
 /sex="hermaphrodite"
 /note="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII
 ; Produced by Rod Wing"
 BASE COUNT 140 a 237 c 120 g 239 t 22 others
 ORIGIN

Query Match 22.6%; Score 354.6; DB 256; Length 758;
 Best Local Similarity 86.7%; Pred. No. 3.4e-79;
 Matches 392; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

QY 894 aggttgatgttgctatgttgcgtggaatgtgtccactactaaagtgatgttctt 953
 |||
 Db 75 AGCTTAGTGTGATGCTGATNGTGGGAATGTGCTCACTAAACGATGCTTCTT 134
 |||

QY 954 catctagagatagcaaatgctcgtctctcactcgttgcctgttgaagcaagatgaca 1013
 |||
 Db 135 CATCTAGAGATAGCAATGCTCTGCTTTCACCTGTGNGTGTGAGCAAGATGACA 194
 |||

QY 1014 ctctgcataagctctcactctgtgtcaacaacctgatgacgtctgcacctgtctcaa 1073
 |||
 Db 195 CTCTGCATCAGCTCTACCTGTGTCAACAACCTGATGACGTGACCGTGTCCAA 254
 |||

QY 1074 ctgctcaactcaaatgttctcgtcgtgtggaacaagaagactgttgcattcgcact 1133
 |||

Db 255 CTGCTCCAACTCCAAATGTTCTGCTGTGGAACAAGAAAGCTTGCCGATTTGACT 314
 |||
 QY 1134 ttatcatctacagatccctgatttggcttctgtgcaagaagagcaagactgact 1193
 |||
 Db 315 TAAGTAATCTACAGATCCCTGATNTGGTCTGGGCAAGAGCANCAAGAACTAACT 374
 |||

QY 1194 tcatgttctccctgcggatgatacagtttgatgattcgcgtctgtgatacatcaag 1253
 |||
 Db 375 CCGATGTGTCCTCGCGGATGATCACTTGATGATTCGCGCTTGATTAACCTCNG 434
 |||

QY 1254 gattcgaagataacggtccaaagtgcgtaccagatttcgacttgcgagatgtgaagto 1313
 |||
 Db 435 GATCCCAAAATNCCGG-CCAGTGCCTCACATATTCGACATTGCGGATGTCNATATC 493
 |||

QY 1314 ttacgtcagctgactctagtttcggttccct 1345
 |||
 Db 494 NNCCCTTTTCTTTTCTTTTCTTTTCTTTCTT 525
 |||

RESULT 7
 H76458 436 bp mRNA EST 05-JAN-1998
 LOCUS H76458
 DEFINITION 18163 Lambda-PRL2 Arabidopsis thaliana cDNA clone 195M2377, mRNA
 sequence.
 ACCESSION H76458
 VERSION H76458.1 GI:1053709
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 436)
 REFERENCE Newman,T., debruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
 L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
 J., and Somerville,C.
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 JOURNAL 95148729
 COMMENT Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tcn@lhm.cl.msu.edu
 Seq primer: 5_prime.
 Location/Qualifiers
 1..436
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone_1ib="Lambda-PRL2"
 /note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not;
 lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques. The vector is BRL's lambda zip-lox. The cDNA
 inserts were directionally cloned with Sal-Not arms using
 oligo dT primed cDNA."

BASE COUNT 109 a 80 c 107 g 124 t 16 others
 ORIGIN

Query Match 19.9%; Score 313; DB 158; Length 436;
 Best Local Similarity 95.1%; Pred. No. 1e-66;
 Matches 331; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

OY 486 gagctagaagaacatctatcgtccctcctcaagaaccacatttgaggaaagtc 545
|||||
Db 1 GAGCTAAGAGAAACACTATCGTCCCTCCCAAGAAACCAACCTTGAGGAAGTCC 60
OY 546 gtattatgtgaatgatctatgtactgtatgtatctctcagtgatgaagaagctta 605
|||||
Db 61 GTATTATGTGAATGATCTATGCTACTGATGATGATCTCTAGATATAGGAAGAGCTTAA 120
OY 606 aggttccctcaagcgaagaataatgaacgtatcgttcgtgagattaacttccctcattg 665
|||||
Db 121 AGGTTCCCTCAACCAAGAAATGAACCTATCGTTCGAGATTATACCTTCTCTTATGG 180
OY 666 aagttctgaacagctctcagagttctcctcaggaagacttaaacatgagcaaga 725
|||||
Db 181 AAGTTCTGAACAGAGCTTCTGAGAGTCTCTCAAGACAGACTATTAACGTATGSCAAGA 240
OY 726 tagctgtctagctctctcgtctgtctcctcaggaagaacgtgtgtgtagg--caag 783
|||||
Db 241 TAGCTGTCTAGCTCTCTCTCTCTCTCTAGGAACACCTCTGTGCTTGAAGGCAAGG 300
OY 784 gaatgggggaatgggctgctgagattagagatcctcattagaagaac 831
|||||
Db 301 NAATGGGGGGAATGGCTTCTGAGCTTAGNGCTTCATTAGGAAC 348

RESULT 8
F14268/c 304 bp mRNA EST 27-JUN-1995
LOCUS F14268
DEFINITION ATTS5190 Gif-Seeda Arabidopsis thaliana cDNA clone YAY910 5', mRNA
sequence.
F14268
F14268.1 GI:871998
VERSION EST
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 304)

REFERENCE
AUTHORS CNRS.
TITLE The Arabidopsis thaliana transcribed genome: the GDR cDNA program
JOURNAL Unpublished (1996)
COMMENT Contact: Berthomieu P., Guerrier D., Giraudat J.
Genetique Moleculaire d'Arabidopsis
ISV - UPR40, CNRS
Avenue de la Terrasse, 1198 Gif-sur-Yvette Cedex, France
Email: Giraudat@cnrs-gif.fr.

FEATURES
source
1. 304
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="YAY910"
/clone_lib="Gif-Seeda"
/note="Vector: Lambda ZAPII non-oriented; Physiological
condition: greenhouse plants. tissue_type: Green siliques"

BASE COUNT 108 a 65 c 47 g 82 t 2 others
ORIGIN

Query Match 19.2%; Score 301; DB 157; Length 304;
Best Local Similarity 99.3%; Pred. No. 1e-65;
Matches 301; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1252 aggttgaagaataagctcaagtgctgtaccagattgacattggagtggtgaaga 1311
|||||
Db 304 AGGATTCGAAGATTAACGGTTCACAGTCCATCAAGATTTCACATTTGGGAGATTTGAAGA 245
OY 1312 tcttcagctagctagcttagttcgttcccttgatcaactgctcctacacatctc 1371
|||||
Db 244 TCTTCAGCTACTGACTTACTTTCGTTTCCTTGATCAACTTGTCTCTATCAACATCTC 185

OY 1372 ttgccattaaaagtttcgaagctcatagatctgtgcttaagttaagtgaagaag 1431
|||||
Db 184 TTGCCATTAAAAGTTTGGAGCTTCATAGAGATCTTCTTAGTAATGTAAGTAGAAG 125
OY 1432 agtttgttttttcgtttatgctttagtaattaaagacataaagtggtgtccg 1491
|||||
Db 124 AGTGTGTTGTTTTCGTTTAGCTTTAGTAATTAATTAACATACAAAGTGtGtGtCCG 65
OY 1492 gattttagatccttaagacataaagccgggttttgaattaggaatcagatttcaat 1551
|||||
Db 64 GATTGTACTAGATCTTAGACATPAACCGCGGTTTGNATTAAGAAATCAGTTTAT 5
OY 1552 gaa 1554
|||||
Db 4 GAA 2

RESULT 9
AV533904/c 487 bp mRNA EST 06-SEP-2000
LOCUS AV533904
DEFINITION AV533904 Arabidopsis thaliana flower buds Columbia Arabidopsis
thaliana cDNA clone FB070B02F 3', mRNA sequence.
AV533904
AV533904.1 GI:8694187
VERSION EST
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 487)

REFERENCE
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1. 487
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="FB070B02F"
/clone_lib="Arabidopsis thaliana flower buds Columbia"
/tissue_type="flower buds"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 196 a 94 c 74 g 123 t
ORIGIN

Query Match 19.2%; Score 301; DB 31; Length 487;
Best Local Similarity 100.0%; Pred. No. 1.2e-65;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1271 ccaagtgctaccagattcgaattcggagatgtgaagatcttcagctagctactc 1330
|||||
Db 487 CCAAGTGGTTACCAAGATTTGCACTTTCGGATGTTGAAGATCTACAGCTAGCTACT 428
OY 1331 agtttggttcttgatcaactgtctcctcctaacaactcttcccataaagtatt 1390
|||||
Db 427 AGTTTGCTTCTTGACCAACTGCTCTATCAACATCTTGCCTCAATAAAGTTT 368
OY 1391 gcaactcatagatctctgtcattgaatgaatgaagaagtgtttgttttcggt 1450
|||||
Db 367 GCAGCTTCATAGAGATCTTGCTTAGTAATGTTAAGGAGAAAGTGTTTGTTCGTT 308
OY 1451 tatgcttagtaattaaagacataaagaagtgtgttccggattgttagtaagatctaa 1510

```

|||||
Db 307 TATGCTTACTAATTAAAGACATACAAAGAGTGTGTTCGGATGTTAGTAAGATCTTAA 248
QY 1511 gacataagccgggttttgcaataggacgcggttttaataagtttagtttagttt 1570
Db 247 GACATTAAGCCGGGCTTTTGCATTAAGATCGAGTTTAATGAAGTTTATGTTATGTTT 188
QY 1571 g 1571
Db 187 g 187

RESULT 10
LOCUS CDS00W2C 231 bp DNA GSS 28-JUN-1999
DEFINITION Arabidopsis thaliana genome survey sequence T7 end of BAC T10P3 of TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.
ACCESSION AL093042 GI:5294196
VERSION AL093042
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 231)
AUTHORS Salanoubat,M., Choisme,N., Artiguenave,F., Brottier,P., Wincker,P., Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 231)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
FEATURES
source
1. .231
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_lib="TAMU"
/clone="T10P3"
/note="end : T7"

BASE COUNT 66 a 63 c 37 g 65 t
ORIGIN

Query Match 14.7%; Score 231; DB 219; Length 231;
Best Local Similarity 100.0%; Pred. No. 5,3e-48;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 662 atggaattctgaaagccttcctgagagttcttctcagacgactactaaactgagggc 721
Db 231 ATGGAATTCTGAAAGCCCTTCGAGAGTTCTTCAGACATCTAAAGCTGAGGCG 172
QY 722 aagatacgtgtcagcgttctcctgctcgttcctcagaaagacgctgtgtgtaagcaa 781
Db 171 AAGATACCTGTGTCAGCTTCTCCTGCTGTTCTCCTAGAGAAAGCCTGTGTTGTTAGGCAA 112
QY 782 aggaatagggggaatgagctgtcgtgagattagatccataaagaactagagacttgg 841
Db 111 AGGAAATGGGGGAATGGGCTGCTGAGATTAAGATCTCTATTGAAGAACTAGAGACTTGG 52
QY 842 ttgggtactttgatactctgaaagacgtgcctaaagcttatgatgttaag 892
Db 51 TTGGGTACTTTGTGATACTCTTGAAGAGCTGTAAAGCTTATGATGCTAAG 1

RESULT 11
LOCUS CDS00TNP 168 bp DNA GSS 28-JUN-1999
DEFINITION Arabidopsis thaliana genome survey sequence T7 end of BAC T4N22 of

```

```

TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.
ACCESSION AL089923
VERSION AL089923.1 GI:5291063
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 168)
AUTHORS Salanoubat,M., Choisme,N., Artiguenave,F., Brottier,P., Wincker,P., Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 168)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
FEATURES
source
1. .168
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_lib="TAMU"
/clone="T4N22"
/note="end : T7"

BASE COUNT 41 a 35 c 38 g 54 t
ORIGIN

Query Match 10.7%; Score 168; DB 219; Length 168;
Best Local Similarity 100.0%; Pred. No. 4,6e-32;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1128 tcgacttactaactacagatccctgatttggttcttcggcagagagcaacaagacc 1187
Db 1 TCGACTTACTAATCTACAGATCCCTGATTTGTTGTCGCGAGAGCAACAAGACC 60
QY 1188 tagacttcgattgttcttcctcgagatcagttgattgatttcgcttcgttgatgaca 1247
Db 61 TAGACTTCGATGTGTTCTCTCGCGATGATCAGTTTGATGATTTGGCTTCGTGATGACA 120
QY 1248 tccaagattcgaagatacagctcccaagtgcgtaccagatttcagact 1295
Db 121 TTCAGGATTCGAAGATACGATCAGATCGTTCACCAAGTTTCGACT 168

RESULT 12
LOCUS Z34128 213 bp mRNA EST 14-JUN-1994
DEFINITION AF053162 Grenoble-B Arabidopsis thaliana cDNA clone GBE349 5', mRNA sequence.
ACCESSION Z34128
VERSION Z34128.1 GI:498484
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 213)
AUTHORS CNRS.
TITLE The Arabidopsis thaliana transcribed genome: the GDB cDNA program
JOURNAL Unpublished (1996)
COMMENT Contact: Mache R., Quigley F., Thomas F., Yu DY.
CNRS URA 1178
Laboratoire de Biologie Moléculaire végétale
B.P. 53X,38041 Grenoble Cedex,France
Email: FOUIGLE@ceicg990.grenet.fr.
FEATURES
source
1. .213
Location/Qualifiers

```

```

/organism="Arabidopsis thaliana"
/strain="ecotype Columbia C24"
/db.xref="taxon:3702"
/clone="G86E349"
/clone_1lb="Grenoble-B"
/note="Vector: Lambda ZAP; tissue=Flower buds of
A.thaliana ecotype columbia C24; clone_library=Grenoble-B;
Cloning vector: Lambda ZAP"

BASE COUNT      47 a      30 c      52 g      83 t      1 others
ORIGIN

Query Match      10.5%; Score 165.4; DB 191; Length 213;
Best Local Similarity 87.4%; Pred. No. 2.2e-31;
Matches 181; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Oy 320 catagttgcgtgagaggttttgagtttgatatacacaagaatttgaacctttaagatgat 379
Db 6 CATAGTTGCGTGAAGAGTTTTCAGTTTATATACAAAAAAGTTTACCTTTTAGAGTGAT 65

Oy 380 ttcttgctcttcgtcttcgtgagttttgagagtggttgatacatggttcagatt 439
Db 66 TTTTGTCTTCTTCCTTTCTTCGTGTTTTCACGAGTCGCTTAAACATGTTTCGATT 125

Oy 440 agaaagacagctctttagtggtgttagtagcagagattaaagaagagctagaagaac 499
Db 126 AGAAAGGACAGACTCTTTAGTGTCTGTTTCGGAAATTTGGAGAGAGCTTAGAGACA 185

Oy 500 actcatgcctccctcccaagaacc 526
Db 186 CTCCTTCGCTCCCTTCCTCCAGAGACC 212

RESULT 13
B20629 763 bp DNA GSS 16-SEP-1997
LOCUS T26B14-Sp6 TMMU Arabidopsis thaliana genomic clone T26B14, DNA
DEFINITION sequence.
ACCESSION B20629
VERSION B20629.1 GI:2395683
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eucots II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 763)
Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
Ecker,J.
BAC End Sequences at ATCC
Unpublished (1997)
Other_GSSs: T26B14-r7
Contact: Ecker J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenome.bio.upenn.edu
Seq primer: Sp6
Class: BAC ends
High quality sequence start: 110
High quality sequence stop: 131.
Location/Qualifiers
1..763
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db.xref="taxon:3702"
/clone="T26B14"
/clone_1lb="TMMU"
/sex="hemaphrodite"
/note="Vector: BelovAC11; Site_1: HindIII; Site_2: HindIII

```

BASE COUNT	143 a	209 c	77 g	317 t	17 others
ORIGIN					
Query Match	9.2%	Score 145.2;	DB 256;	Length 763;	
Best Local Similarity	79.9%;	Pred. No. 3.9e-26;			
Matches 187;	Conservative 0;	Mismatches 45;	Indels 2;	Gaps 2;	
Oy	903	ttgatgtcatgtgtgct-ggaatgtgtccactactaactgtatgtttcttcacgtag	961		
Db	81	TTGAAGCATATNGTNGCTGGGAAATCGCCACCTACTAAACGTGATTTCTTCACATCAG	140		
Oy	962	actagccaatgtctcgtctgtcttcacacctgttctcgtgttgagcaagaatgacactctgca	1021		
Db	141	AGTACCCACCGCTCTCGTNCNTACACTGTNGTNCCTGTNGAGCAAGATGACACTTCTGCA	200		
Oy	1022	tcagctctcaactgtgttgcacaacacctgtagaagctctgcacctgtgtctcaactgtcca	1081		
Db	201	TCAGCTCTCACTTGTGTGTCAACACCCCTGATGACGTCTGCAAGNNG-NCCTGCTGNTTCA	259		
Oy	1082	actccaatgttccctgcgtggtggaagaagaagatgtgttcagatttcgacttt	1135		
Db	260	ATTCCTATTTTTCTGCTGTGGCGTAAATAGCAGCACCTTGTTATGTATTATTTT	313		
RESULT 14					
BE524424					
LOCUS	BE524424	410 bp	mRNA	EST	19-MAR-2001
DEFINITION	M49G95TM Arabidopsis developing seed Arabidopsis thaliana cDNA				
ACCESSION	BE524424				
VERSION	BE524424.1	GI:9782402			
KEYWORDS	EST.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 410)				
AUTHORS	White,J.A., Todd,J., Newnan,T., Focks,N., Gire,J.T., Martinez de Ibarra,O., Javorski,J.G., Ohlrogge,J. and Benning,C.				
TITLE	A new set of Arabidopsis expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil				
JOURNAL	Plant Physiol. 124 (4), 1582-1594 (2000)				
MEDLINE	20567808				
COMMENT	Contact: Benning, C Dept. of Biochemistry & Molecular Biology Michigan State University 224 Biochemistry, Michigan State University, East Lansing, MI 48824 , USA Tel: 517 355 1609 Fax: 517 353 9334 Email: benning@msu.edu				
FEATURES	Location/Qualifiers				
source	1. 410 /organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /clone="M49G9" /clone_id="Arabidopsis developing seed" /tissue_type="seed" /dev_stage="5-13 days after flowering" /lab_host="E. coli" /note="Organ: Developing seed; Vector: pBluescript SK-; Site_1: EcorI, Site_2: XhoII"				
BASE COUNT	97 a	83 c	88 g	142 t	
ORIGIN					

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2001, 17:00:46 ; Search time 7917.7 Seconds
(without alignments)
1362.227 Million cell updates/sec

Title: US-09-394-519-123

Perfect score: 1141
Sequence: 1 aggaataaataacatgaty.....gtacaaaaaaaaaaaaa 1141

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq-length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:**

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_estl48:*
118: gb_estl49:*
119: gb_estl50:*
120: gb_estl51:*
121: gb_estl52:*
122: gb_estl53:*
123: gb_estl54:*
124: gb_estl55:*
125: gb_estl56:*
126: gb_estl57:*
127: gb_estl58:*
128: gb_estl59:*
129: gb_estl60:*
130: gb_estl61:*
131: gb_estl62:*
132: gb_estl63:*
133: gb_estl64:*
134: gb_estl65:*
135: gb_estl66:*
136: gb_estl67:*
137: gb_estl68:*
138: gb_estl69:*
139: gb_estl70:*
140: gb_estl71:*
141: gb_estl72:*
142: gb_estl73:*
143: gb_estl74:*
144: gb_estl75:*
145: gb_estl76:*
146: gb_estl77:*
147: gb_estl78:*
148: gb_estl79:*
149: gb_estl80:*
150: gb_estl81:*
151: gb_estl82:*
152: gb_estl83:*
153: gb_estl84:*
154: gb_estl85:*
155: gb_estl86:*
156: gb_estl87:*
157: gb_estl88:*
158: gb_estl89:*
159: gb_estl90:*
160: gb_estl91:*
161: gb_estl92:*
162: gb_estl93:*
163: gb_estl94:*
164: gb_estl95:*
165: gb_estl96:*
166: gb_estl97:*
167: gb_estl98:*
168: gb_estl99:*
169: gb_estl00:*
170: gb_estl01:*
171: gb_estl02:*
172: gb_estl03:*
173: gb_estl04:*
174: gb_estl05:*
175: gb_estl06:*
176: gb_estl07:*
177: gb_estl08:*
178: gb_estl09:*
179: gb_estl10:*
180: gb_estl11:*
181: gb_estl12:*
182: gb_estl13:*
183: gb_estl14:*
184: gb_estl15:*
185: gb_estl16:*
186: gb_estl17:*
187: gb_estl18:*
188: gb_estl19:*
189: gb_estl20:*

190: gb_estl21:*
191: gb_estl22:*
192: gb_estl23:*
193: gb_estl24:*
194: gb_estl25:*
195: gb_estl26:*
196: gb_estl27:*
197: gb_estl28:*
198: gb_estl29:*
199: gb_estl30:*
200: gb_estl31:*
201: gb_estl32:*
202: gb_estl33:*
203: gb_estl34:*
204: gb_estl35:*
205: gb_estl36:*
206: gb_estl37:*
207: gb_estl38:*
208: gb_estl39:*
209: gb_estl40:*
210: gb_estl41:*
211: gb_estl42:*
212: gb_estl43:*
213: gb_estl44:*
214: gb_estl45:*
215: gb_estl46:*
216: gb_estl47:*
217: gb_estl48:*
218: gb_estl49:*
219: gb_estl50:*
220: gb_estl51:*
221: gb_estl52:*
222: gb_estl53:*
223: gb_estl54:*
224: gb_estl55:*
225: gb_estl56:*
226: gb_estl57:*
227: gb_estl58:*
228: gb_estl59:*
229: gb_estl60:*
230: gb_estl61:*
231: gb_estl62:*
232: gb_estl63:*
233: gb_estl64:*
234: gb_estl65:*
235: gb_estl66:*
236: gb_estl67:*
237: gb_estl68:*
238: gb_estl69:*
239: gb_estl70:*
240: gb_estl71:*
241: gb_estl72:*
242: gb_estl73:*
243: gb_estl74:*
244: gb_estl75:*
245: gb_estl76:*
246: gb_estl77:*
247: gb_estl78:*
248: gb_estl79:*
249: gb_estl80:*
250: gb_estl81:*
251: gb_estl82:*
252: gb_estl83:*
253: gb_estl84:*
254: gb_estl85:*
255: gb_estl86:*
256: gb_estl87:*
257: gb_estl88:*
258: gb_estl89:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

JOURNAL MEDLINE COMMENT	size-selected cDNA libraries DNA Res. 7, 175-180 (2000) 20363093 Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 153-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.			
FEATURES	Source			
	1. 443 /organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /clone="S017jell1P" /clone_id="Arabidopsis thaliana green siliques Columbia" /tissue_type="green siliques" /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"			
BASE COUNT	136 a	107 c	89 g	111 t
ORIGIN				
Query Match	38.8%; Score 443; DB 31; Length 443;			
Best Local Similarity	100.0%; Pred. No. 6.9e-110;			
Matches 443; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;
QY 681	tgcgaactcttgaaacggtgctgtatgagaagctaaacggaagaaacggaagcttcaagaa	740		
DB 443	TGCGAATATCTTGAAACGGGTGGTAGAGAGCTTAACGGAAGAACCGGAGACTTCAGAAA	384		
QY 741	gagcctatgagcttcgaactccaagctgtctccacaattctacggtcagatgctcca	800		
DB 383	GAGGCTATGGAGCTTCCAGACTCTCAAGCTGTCTCCACAATTCTACGGTCAAGTACTCCA	324		
QY 801	ccaactcacatcatcatgctgtccttcgtgagagcgtygtgggtggccatcatcagaa	860		
DB 323	CCAACTACACTCATCATGATGTGCTTCGTCGAGCAGCGTGGGGCCCATCATCATCGAAC	264		
QY 861	catcaccaaatcaacagagccggtttctatacaatccgtyggtgtgctgtgctgaagt	920		
DB 263	CATCACCACAATTCACAGAGCCCGTTTCTATCAATCCGTGGGTTGCTTGCTGCTCAGGTG	204		
QY 921	gtcatagtgctgaattttgaagccttggtccaagatgtaattttatttttagtggt	980		
DB 203	GCTCATGGGCTGAATTTTGAAGCCCTGGCTGCACAGCATGTATTTTATTTTAACTGAGTG	144		
QY 981	ggaagaagtggtttgtgatttttcgtatcgttatatactatgctatctgtgtggagct	1040		
DB 143	GGAAAAGGGCTGTTTGGATTTTTCGTTATGCTATATAGTCTATCTGTGGGTCATT	84		
QY 1041	gtaatttgatgatctgagccttcatactgaactagtcctatglatgatgacaacttaaaa	1100		
DB 83	GTAATTTTGGATGATGTTGGCCCTTCATCATGACTAGTCATGATGATGATGATCAACCTTA	24		
QY 1101	gatttaaatgacaaaattagt	1123		
DB 23	GATTTAATTTAGCAAAATTAAGT	1		
RESULT 2				
LOCUS	AT100243	545 bp	mRNA	EST 21-AUG-1998
DEFINITION	34618 lambda-PRU2 Arabidopsis thaliana cDNA clone 110C16P 3', mRNA			
ACCESSION	AT100243			
VERSION	AT100243.1 GI:3450204			
KEYWORDS	EST.			
SOURCE	thale cress.			
ORGANISM	Arabidopsis thaliana Eukaryotes: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eucotids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 545)			

AUTHORS		TITLE		JOURNAL		COMMENT			
Newman, T., deBrujin, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlirrogge, J., Ralkehl, N., Somerville, S., Thomashow, M., Retzel, E., and Somerville, C.		Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones		Plant Physiol. 106, 1241-1255 (1994)		95148729			
Contact: Thomas Newman		MSU-DOE Plant Research Laboratory		Michigan State University		Lansing, MI			
Tel: 517-353-0854		Fax: 517-353-9168		Email: 22313ycn@hm.cl.msu.edu		The sequence entry for this EST has been reverse complimented and is being submitted in the sense orientation.			
Seq primer: M13-Universal.		Location/Qualifiers		1. 545					
/organism="Arabidopsis thaliana"		/strain="var columbia"		/db_xref="taxon:3702"					
/clone="110C16XP"		/clone_id="Lambda-PRL2"		/note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark - rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRU's lambda zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."					
BASE COUNT		129 a 116 c 131 g 163 t 6 others		ORIGIN					
Query Match		36.6%; Score 418; DB 16; Length 545;		Best Local Similarity		96.4%; Pred. No. 4.7e-103;			
Matches 449; Conservative 0; Mismatches 12; Indels 5; Gaps 2;									
0Y	661	taaaccaaaaggagtagatgctgaatacttgaaaaggtgctgaagaagcacaaggaag	720						
DB	85	TTAAANNAACGGAGGTAGATT---CGACTGAACGGGTCTAGAGAAGCTAACGGAAG	140						
0Y	721	agaacccgaagactcaagaagaagagctatagagcttcgaactccaagctgtctccacaat	780						
DB	141	AGACC CGGAGACTCAG---AAGAGGCTATGAGACTTCGAACCTCTCAAGCTGTCCACAAT	199						
0Y	761	tctaaggtcagatgctccaccaactacaatcatatgtctcttcgvgagaggtgtg	840						
DB	200	TCTAGCGTCAGATGACTCCACCACATCACTCATGTCTTCCTCGCGGCGCTGTGG	259						
0Y	841	gtggccatcatcatcgcgaaccacacacacaatcacaggccgcttcatacactcggtg	900						
DB	260	GTGGCCATCATCATCATCGAACCATCAACACACAGAGCCGCTTCTATCATCCGTGG	319						
0Y	901	ttgcttctgtgctcagtgagctcatagggctgaatttgaagccttggtccagatcgt	960						
DB	320	TTGCTTGTGCTGTGTCAGGTGCTCATGTGGCTGAATTTGGAACCTTGGCTCCACATCGT	379						
0Y	961	gatttttatttaattggtgggaaaagggtgtttgtgatttttgatatgatatag	1020						
DB	380	GATTTTATTTTATTTAATGTGGGGAAGGCTTTTGTGTAATTTTGTATCGTATATAG	439						
0Y	1021	tctatctgtgtgggtcatgtcaattttgagatgtgcccctcctaactacagccat	1080						
DB	440	TCTATCTGTGTGGGTCAATGTAAATTTTGGATGATTTGGCTTCTCATGAACTAGTCCAT	499						
0Y	1081	gtatgagcaccttaaaaagatttaattgaataaattagtaaatagttac	1126						

DB	500	GTATGATGCACCTTAAAAAGATTAAATTACGCAAAATTAAGTTAC	545
RESULT	3		
LOCUS	AV556109/C		
DEFINITION	AV556109 Arabidopsis thaliana green siliques Columbia Arabidopsis thaliana cDNA clone S0034902F 3', mRNA sequence.		
ACCESSION	AV556109		
VERSION	AV556109.1		
KEYWORDS	GI:8727524		
SOURCE	EST.		
ORGANISM	thale cress.		
SOURCE	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 375)		
AUTHORS	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.		
TITLE	A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized site-selected cDNA libraries		
JOURNAL	DNA Res. 7, 175-180 (2000)		
MEDLINE	20363093		
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yama 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/		
FEATURES	Location/Qualifiers		
source	1..375		
	/organism="Arabidopsis thaliana"		
	/strain="Columbia"		
	/db_xref="taxon:3702"		
	/clone="S0034902F"		
	/clone_11b="Arabidopsis thaliana green siliques Columbia"		
	/tissue_type="green siliques"		
	/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"		
BASE COUNT	126 a 84 c 78 g 87 t		
ORIGIN			
Query Match	32.7%; Score 373.4; DB 31; Length 375;		
Best Local Similarity	99.7%; Pred. No. 66-91; Indels 1; Gaps 0		
Matches 374:	Conservative 0; Mismatches 1; Indels 0; Gaps 0		
QY	752 gcttcgaactctcaagctgctctcaacaattctacggtcagatgactccaccaactaact	811	
DB	375 GCTTGAACCTCAAGCTCTCTCCCAATTTCAAGCTCAGATGACTCCACCAACTACACT	316	
QY	812 catctgtgctcttggtgagcgt	871	
DB	315 CATCATGTCCTCTGCTGTGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	256	
QY	872 tcaagagcccggttctcatcaacacgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	931	
DB	255 TCACAGGCCCGTTTTCATCAATCCGAGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	196	
QY	932 gaatttgaagccctgcgtccacagatcgtgatttttattatttaagtggtgtggaaggtg	991	
DB	195 GAATTTTGAAGCCTTCGCGTCCACGACGCGGATTTTATTTTAAGTGTGGGAAAGGGTG	136	
QY	992 ttgttgattttctgtatcgtataatagctatcctatcgtgtgtgtgtgtgtgtgtgtgt	1051	
DB	135 TTTTGTATTTTTCGTTATTCGTTATATAGTCATCTCTGTGTGTGTGTGTGTGTGTGT	76	
QY	1052 tgaatggcccttcacgaactagctcctatgatatgacatgacacacttaaaagatttaata	1111	
DB	75 TGATGGCCCTTCATGAAATAGTCTATGATGATGATGATGATGATGATGATGATGATG	16	
QY	1112 gcaaaatagttac 1126		
DB	15 GCAAAATATTAGTTAC 1		

RESULT	4	TA2030	475 bp	mRNA	EST	07-JAN-1998
LOCUS	5293	Lambda-PRL2	Arabidopsis thaliana	CDNA clone	110C16T7	mRNA sequence.
DEFINITION	TA2030	GI:932917				
ACCESSION	1	AGGAAATATAAAAACATGATGATGGCCAAAGATCTAGTTTGAGCCTAAGCTTAAGGCT				
VERSION	61	TTTCACAAATACACATCTCTTCAGATGAAATCGAATCTTAACCTTCATTATCAACA				
KEYWORDS	121	attcacaataacgaatcctcttagatgaatcgaatcccaactcttattataaaca				
SOURCE	121	attcacaagaagctttccatcaacaggttaactcgcgaggaagaacaagaggttttcgtaccaa				
ORGANISM	121	ATCTCCAGAGACTCCCAATGGAACCAAAACATTTCGATCTCAATCAATCTTCGCAAGATAG				
REFERENCE	1	aggaataataaacaatgatgattggcacaagaagatcagtttgagcctaagctagaggt				
AUTHORS	1	acgugaagaagttttccatcaacaggttaactcgcgaggaagaacaagaggttttcgtaccaa				
TITLE	181	ACGTGAACAGTTTTCATCAATCAACGGTTAACTCGAGGAAGACACAGAGATTTTCGTACCAA				
JOURNAL	181					
MEDLINE	181					
COMMENT	181					
FEATURES	181					
SOURCE	181					

QY	241	acagacgacatcccaagccacattagccggaaagagaaagtggagaaagaaagaaatctccggaa	300
Db	241	ACAGTACGATCTCAAGCCACCTTACCGGAAAGAGAGTGAAGAGAACGATCTCCGGAA	300
QY	301	ccggatcttgctccggcgacgatacgaacgagatcactccgatacgaagg--lactaacg	358
Db	301	CCGGGCTTGCGCTCCGGCAGCAGTACAGCGGAGTACTCCGATGACGAGGGTACTTCACG	360
QY	359	tggaactccagat----gaagaagaagcggggcgaaactctgaggaagaagctcaggt	414
Db	361	TGGAACTTCAGATTGAGAGAGAACGGAGGCGGAAACGTGCGNMGGAAGAAANTTANGT	420
QY	415	tatcaaaagatcagctgctcttctccgaagagacttcaagaacaacacactc	469
Db	421	TNTTCAAGATCGNCCCTCTTTTTCG-GGGGNTTNCANGNMCANANNTTT	474
RESULT	5		
LOCUS	AV556091/c		
DEFINITION	AV556091 Arabidopsis thaliana green siliques Columbia Arabidopsis		
ACCESSION	AV556091		
VERSION	AV556091.1		
KEYWORDS	EST.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
TITLE	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
JOURNAL	1 (bases 1 to 372)		
MEDLINE	Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.		
COMMENT	A large scale analysis of cDNA in Arabidopsis thaliana: Generation		
	of 12,028 non-redundant expressed sequence tags from normalized and		
	size-selected cDNA libraries		
	DNA Res. 7, 175-180 (2000)		
	Contact: Erika Asamizu		
	The First Laboratory for Plant Gene Research		
	Kazusa DNA Research Institute		
	Yana 1532-3, Kisarazu, Chiba 292-0812, Japan		
	Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.		
FEATURES	location/Qualifiers		
source	1..372		
	/organism="Arabidopsis thaliana"		
	/strain="Columbia"		
	/db_xref="taxon:3702"		
	/clone="S0034C01P"		
	/clone_lib="Arabidopsis thaliana green siliques Columbia"		
	/tissue_type="green siliques"		
	/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:		
	XhoI"		
BASE COUNT	122 a 91 c 78 g 81 t		
ORIGIN			
Query Match	31.2%; Score 356; DB 31; Length 372;		
Best Local Similarity	97.3%; Pred. No. 3.3e-86;		
Matches 362; Conservative 0; Mismatches 10; Indels 0; Gaps 0;			
QY	731	acttaagaagaagatagagagcttcgaagctcaagctcaagctctcacaactctacgta	790
Db	372	ACTTCAAGAAAAAGCTATGAGAGTTTGAGACCTCTCAAGCTGCTCCACAGGTTTACGGGTCA	313
QY	791	gatgactccaccaatacactcatcatcatatgtgtctctcgtgtagcgctgtggtgagccatc	850
Db	312	GATGACTCCACCAACTATACATCATCATGTGCTTCGTCGACGAGCGTGTGGGTGGCCCATC	253
QY	851	atcatcgacacatcaaccacatcaacgaagccggtttctatcaatccatcggtgtgttctgtc	910
Db	252	ATCATCGAACCATCCCCACATACAGGCGCGGTTTCTATCATCATCGGTGGCTTGCTTGTC	193

QY 911 tggcagtgctcatgagctgaatttgaagccttcgctccacagctgattttat 970
 |||||||
 Db 192 tggtagtgctcatgagctgaatttgaagccttcgctccacagctgattttat 133
 |||||||
 QY 971 tttagtgctgaggaagagctgcttgcgtatcttcgtatcgtatatagtctatctgtg 1030
 |||||||
 Db 132 tttagtgctgaggaagagctgcttgcgtatcttcgtatcgtatatagtctatctgtg 73
 |||||||
 QY 1031 tgggctcatgctgaatttgaagcttccttcctcctgaagctcctatgctatgctga 1090
 |||||||
 Db 72 tgggctcatgctgaatttgaagcttccttcctcctgaagctcctatgctatgctga 13
 |||||||
 QY 1091 accttaaaaaa 1102
 |||||||
 Db 12 accttaaaaaa 1
 |||||||

RESULT 6
 AI994446 526 bp mRNA EST 08-SEP-1999
 LOCUS 701498340 A. thaliana, Ohio State clone set Arabidopsis thaliana
 DEFINITION CDNA clone 701498340, mRNA sequence.
 ACCESSION AI994446
 VERSION AI994446.1 GI:5841351
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 526)
 AUTHORS Chen, J., Momiyama, M., Chan, E., Mooney, M., Carreon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobriza, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and Hanson, D.
 TITLE Arabidopsis thaliana Gene Expression Microarray
 JOURNAL Unpublished (1999)
 COMMENT Contact: David Smoller, Ph.D.
 Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.
 4633 World Parkway Circle, St. Louis, MO 63134, USA
 Tel: 877-577-2733
 Fax: 314-427-3324
 Email: service@genomesystems.com.

FEATURES
 source
 1. 526
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="701498340"
 /note="CDNA library was made from selected clones from the Arabidopsis thaliana Ohio State clone set."
 BASE COUNT 170 a 122 c 103 g 130 t 1 others
 ORIGIN

Query Match 29.7%; Score 338.4; DB 104; Length 526;
 Best Local Similarity 96.5%; Pred. No. 2.2e-81;
 Matches 356; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1 aggaataaataacatgatgatgaggaagaagatcgttagtgagcctaagcttaggt 60
 |||||||
 Db 148 aggaataaataacatgatgatgaggaagaagatcgttagtgagcctaagcttaggt 207
 |||||||
 QY 61 ttccacaaatcacacatcctcttcagatgaatcgaatcctaactcttcatatcaaaa 120
 |||||||
 Db 208 ttccacaaatcacacatcctcttcagatgaatcgaatcctaactcttcatatcaaaa 267
 |||||||
 QY 121 atctcagagactccatggaacaaacatctcgatctctcatcagatcttcgcaagatg 180
 |||||||
 Db 268 atctcagagactccatggaacaaacatctcgatctctcatcagatcttcgcaagatg 327
 |||||||

QY 181 acgtgaacagctttccatcaacgatttaactgaggaagacacagaggttcgtccaaa 240
 |||||||
 Db 328 acgtgaacagctttccatcaacgatttaactgaggaagacacagaggttcgtccaaa 387
 |||||||
 QY 241 acagtaacagcttcac 300
 |||||||
 Db 388 acagtaacagcttcac 447
 |||||||
 QY 301 ccggagctgctc-tccggagacacacacacacacacacacacacacacacacacac 359
 |||||||
 Db 448 acggagctgctc-tccggagacacacacacacacacacacacacacacacacacac 507
 |||||||

RESULT 7
 H76307 390 bp mRNA EST 30-DEC-1997
 LOCUS 18012 Lambda-PRL2 Arabidopsis thaliana CDNA clone 200B977, mRNA
 DEFINITION sequence.
 ACCESSION H76307
 VERSION H76307.1 GI:1053558
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 390)
 AUTHORS Newman, T., deBrujin, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, J. and Somerville, C.
 TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis CDNA clones
 JOURNAL Plant Physiol. 106, 1241-1255 (1994)
 MEDLINE 95148729
 COMMENT Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E. Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tne@msu.edu
 Seq primer: 17 dye primer.

FEATURES
 source
 1. 390
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone="200B977"
 /clone="Lambda-PRL2"
 /note="Vector: lambda ZAP-Lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a CDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRU's lambda ZAP-Lox. The CDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."
 BASE COUNT 94 a 79 c 95 g 107 t 15 others
 ORIGIN

Query Match 27.4%; Score 312.4; DB 158; Length 390;
 Best Local Similarity 90.4%; Pred. No. 2.5e-74;
 Matches 347; Conservative 0; Mismatches 35; Indels 2; Gaps 2;

Oy	654	accgaattaaagcaacgagtagattgycgaatacttgaacagggtgcgtagaagacta	713
Db	1	ACCAAGTTAAAGCAACCGAGGTTACATTGGCAATATCTTGAAACGGTGGCTAGAGAACTA	60
Oy	714	acggaagagacccgagagcttcaagaagagctatgtagcttcgagacttcaagttgct	773
Db	61	ACGGAAGAAACCGGAGACTTCAGAAAGGCGTTATGAGAGCTTCGAACCTCGAAGCTGTCT	120
Oy	774	ccgaattctacggtcagatgagctccccaactacatcatcatatggttccttgtagag	833
Db	121	CCACAATTNTNGGTCACATGATGACTCCACCACACTACATCATCATCTTNNCTTCGTGCAG	180
Oy	834	cgtagggtagcccatcatcatcagaccatcaaccacaatccaggccggtttcttcaat	893
Db	181	CNTNNGGTGGCCCATCATCATCGAACCATCATCCACAAATCACAGGCCGCGTTTNTATCAAT	240
Oy	894	ccgtgagttgcttgctgctgcgtcagtagtgctatggtcgctga-a-tttgaagcttgctcc	952
Db	241	CCGTGGGTGTGCTGTGCTGCTGCTAGGTGGCTCATGGGCTGAATTTTAAACGTTTNGTCC	300
Oy	953	acgaatcgtag-a-ttttattttagtgctgctgagaaagggtgttttgtaattttcgtatc	1011
Db	301	ACGNCGAAATTTTNTTNTTAAAGTGTTGGGAAAGAGGTTTGGGATTTTCGGGNATC	360
Oy	1012	gttatatagctatctgtgtgagg	1035
Db	361	GNAANATGGCTTNTCTTTGTGGG	384
RESULT	8		
LOCUS	AV556640	276 bp	mRNA
DEFINITION	AV556640 Arabidopsis thaliana green siliques Columbia Arabidopsis thaliana CDNA clone SQ048f05f 3', mRNA sequence;	EST	06-SEP-2000
ACCESSION	AV556640		
VERSION	AV556640.1	GI:8728055	
KEYWORDS	EST.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustos II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 276)		
TITLE	Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.		
	A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence:tags from normalized and size-selected cDNA libraries		
JOURNAL	DNA Res	7, 175-180	(2000)
MEDLINE	20363093		
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/		
FEATURES	Location/Qualifiers		
Source	1..276		
	/organism="Arabidopsis thaliana"		
	/strain="Columbia"		
	/db_xref="taxon:3702"		
	/clone="SQ048f05f"		
	/clone_1lb="Arabidopsis thaliana green siliques Columbia"		
	/tissue_type="green siliques"		
	/note="Vector: pbluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"		
BASE COUNT	97 a	68 c	50 g
ORIGIN			61 t
Query Match	23.6%;	Score 269.6;	DB 31; Length 276;
Best Local Similarity	98.6%;	Pred. No. 1e-62;	
Matches	272; Conservative	0; Mismatches	4; Indels
		0; Gaps	0;
836	tgtagggtagccatcatcatcgaaccatcacacaatcacagggccggtttctatcatcc	895	

[illegible]

RESULT	9
LOCUS	AI0099721
DEFINITION	Arabidopsis thaliana cDNA clone 120N14XP 3', mRNA
ACCESSION	AI0099721
VERSION	AI0099721
KEYWORDS	EST.
SOURCE	GI:3449460
ORGANISM	thale cress. Arabidopsis thaliana
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucosids II; Brassicales; Brassicaceae; Arabidops.
AUTHORS	1 (bases 1 to 389) Newman,T., deBorja,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel,E., and Somerville,C.
TITLE	Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994)
JOURNAL	95148729
MEDLINE	Contact: Thomas Newman MSU-DOE Plant Research Laboratory Michigan State University MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI Tel.: 517-353-0854 Fax: 517-353-9168 Email: 22313@cneidm.cl.msu.edu
COMMENT	The sequence entry for this EST has been reverse complemented and is being submitted in the sense orientation. Seq primer: M13-21.
FEATURES	Location/Qualifiers
SOURCE	1..389 /organism="Arabidopsis thaliana" /strain="var columbia" /db_xref="taxon:3702" /clone="120N14XP" /clone_id="Lambda-PRL2" /note="Vector: lambda zip-lox; site_1: Sal; site_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr airtel, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA."
BASE COUNT	88 a 83 c 82 g 130 t 6 others
ORIGIN	


```

/clone="T27C23"
/clone_id="RAMU"
/sex="hermaphrodite"
/note="vector: Belosacit; Site_1: HindIII; Site_2: HindIII
; Produced by Rod Wing"
172 a 102 c 120 g 140 t

```

Query Match	14.7%	Score 167.6;	DB 256;	Length 534;
Best Local Similarity	70.1%	Pred. NO. 7.1e-35;		
Matches 269;	Conservative 0;	Mismatches 109;	Indels 6;	Gaps 13;

QY	477	aaacagaagctgcttctggtctgaagcgtgaacttgacggccaagtggaagtctgg	536
Db	152	AAACAGAGAGCTGGCTTTGGCTAAGAACTAGCTTTAACGACCAAGACAGTGGAAGTGTGG	211

Qy	537	ttcccaaacagaaagactgagtgagaa-----tttctgtcttacaacccaataattgtataaa	552
Db	212	ttcccaaaaacagaaagagcacaagtgtaagaaactgaaagtcgtcatattattttttacacattttactgt	271

QY	593	ttacattataccccctgtgtgtagaagttttac-ttcaagaglcagatttgatttca	651
Db	272	GACCAATTGGCCCTAAAGAGCATGTATTTCACATCAACACCATTGTTGTCCTAATA	331

Oy 652 gaaccäaglttaagaacaaacggaagttagattltagatacttgaacacgtgcgtatagaagc 711
| | | | | | | | | | | | | | | | | | | | | |
Db 332 ggcacaaactttaagcacagacccaagatggatttcacagtatatttgaaagaacgtcttcagagaat 391
| | | | | | | | | | | | | | | | | | | | | |

OY 712 taacgaagaagaaaccggagacttcagaaaagagctatgagctcgcgaactctcaagctgt 771
|||||
DB 392 TAACGGAAGAGAATTCGGCGCTTGAGAAAAGA -TCAGCGCAACTAAGACCATTAATAAGCTTTT 450
|||||

0y 772 ctccacaaattctcagctcagatctactccaccacaaactacatcatcatgtctcttcgtcg 831
 : :
 db 451 cctccgcctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 510
 : :

QY 832 agcgtgtgggtgccccatcatcat 855
1 ||||| 1 ||||| 1 |||||
511 TTTTTCCTCTCTCTCTCTCTCTT 524

RESULT 12

LOCUS	531 bp	EST	06-SEP-2000
AV551157	Arabidopsis	thaliana roots	Columbia Arabidopsis thaliana
AV551157	CDNA clone R212a06r 5', mRNA sequence.		

```

VERSION      AV551157.1  GT:8722570
KEYWORDS
SOURCE       EST.
             thale cress.

```

ORGANISM
Eubryopsis, Chlamydomonas, Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, Rosidae, eurosid II, Brassicales, Brassicaceae, Arabidopsis.

REFERENCE	1
(cases 1 to 31)	
AUTHORS	Asamitsu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE	A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and

JOURNAL
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093
COMMENT
Contact: Erika Asamien

The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 252-0812, Japan
Tel: 0477-437-0123 Fax: 0477-437-0471

FEATURES	Location/Qualifiers
SOURCE	1. 531
	/organism="Arabidopsis thaliana"

```
/db_xref="taxon:3702"
/clone="R2122a06r"
```

```

/clone.lib="Arabidopsis thaliana roots Columbia"
/tissue_type="roots"
/node="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
163 a 115 c 110 g 143 t

```

Query Match	14.38;	Score 162.8;	DB 31;	Length 531;
Best Local Similarity	-68.88;	Pred. No. 1.4e-33;		
Matches 278;	Conservative 0;	Mismatches 102;	Indels 24;	Gaps 3;

QY 9 aaaaacgtatbatggccaagaagatctagtggctgaagcttaggatttcaca 68
| | | | | | | | | | | | | | | |
Db 133 AATATCATGATCATTGGGTAAAGACGATTGGCTTTAACTTTCAGCTTGCGATTGCACAA 192

Qy 69 atcacacattccctccttgtagtgcgaactgtaacccttcataatcaaacactctcgag 128
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 AACCC--ATCTTCCTCCACTAATCTTTAAACCACTTCTCACCANAATGTCANNTCTCGAG 249

Qy	129	agactcccatggaaccaaatcattgcattctacatcaga-----tctt	170
Db	250	atggtttccatgacaccaaacccttggttcttccctcagagtcaacaaagcaacagttttt	309

171 cgcaagatagacgltgaacagttlccatcaacggttaactcgaggaagacacaggagtt 230
Oy
310 aggaataatcgacctgaacagcttggccaacacacgctggattttggaagagacagacaggactt 369
Db

Qy 231 tcgtcaccacaacagtlacgatactcgaagccatlagcgggaagagaagtgtagagagaaga 230

Dh 370 TCGTCTCCCAACAGTACGATCTCGAGACACAGTASGTGGAAAAGGAG--GAGTACTGAA 426

[illegible]

QY 351 tactcacgtgaaacctcagatgaagaagaacgcggcgcaaac 394
| | | | | | | | | | | | | | |
DB 487 tctcttacttcgcatccctccattccatgacgataaccgacaac 530
| | | | | | | | | | | | | | |

RESULT 13

LOCUS	569 bp	mRNA	EST	07-SEP-2000
AV5533327				
AV5533327	Arabidopsis	thaliana roots	Columbia	Arabidopsis thaliana
CDNA	Clone RZ60803R 5',	thaliana		
	sequence.			

VERSION	AV553327.1	GI:8724740
KEYWORDS		
EST		
SOURCE	thale cress.	

CONGENERISM

Eukaryota: Eukariotae
Eukaryota: Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;
Eukaryota: Magnoliophyta; eudicotyledons: core eudicots;
Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
1 (uses 1 to 203)
AUTHORS
Asamitsu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and

JOURNAL
MEDLINE
COMMENT
Contact: Erika Asamizu
DNA Res. 7, 175-180 (2000)
size-selected cDNA libraries
20363093

The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 299-0812, Japan
E-mail: kazusa@kazusa.or.jp or kazusa@plant/

```

FEATURES
source      Location/Qualifiers
1..569     /organism="Arabidopsis thaliana"

```

```
/db_xref="taxon:3702"  
/clone="RZ60a03R"  
/clone_1lb="Arabidopsis thaliana roots Columbia"
```


JOURNAL
COMMENT

of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCCTATAGG
High quality sequence stop: 746.

FEATURES

location/Qualifiers
1..860

/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0011F23f"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
/lab_host="E. coli"
/note="Vector: PBK-CMV; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 257 a 190 c 222 g 186 t 5 others
ORIGIN

Query Match 12.3%; Score 140.8; DB 153; Length 860;

Best Local Similarity 66.0%; Pred. No. 1.6e-27;

Matches 268; Conservative 0; Mismatches 102; Indels 36; Gaps 3;

```
OY 152 cgatcctacatcagatcttcgaaagatagacgtaaacagtttccatcaacggttaactg 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 CGAGACTAGATCATTTCTTAGAGAGATCGACGTGAACAGATTGCCATCTACGTCGATTG 66

OY 212 cgaagaagacacagagatttgcacaaagatagatctcaagcaaccattagcggaa 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 CGAAGAAGAAAGCTGAGTTTCATCTCCGACACGACGATATCGAG---TGTAGTGGGAA 123

OY 272 gagaagtgagaagaagaatctcgaaccgcgttgcctcgcgacgatacagacga 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 AAGGAGCGAAGGAGGACACTATATGA-----GATGA 156

OY 332 gatcactcggatcgaagggtactcaagtggaacctcagatgaagaagaagcggggcga 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 157 GCTTGATATTGAAGAGCTGTCTCGTGGCATCATGATGAGGAAGAT-----GGTGA 210

OY 392 aacgtcgaagaagaagctcaggtatatacaaaagatcagtcgtcttctcgaagaagactt 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 CGTTTCGAGAAAAAACTTAAGCTTTTCGAAGGATCAGTCTGTATTCTTGAAGAAAGCTT 270

OY 452 caaagaacacacactctcaatcccaaaagaaagcttagcttgctaaagaagctgaact 511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 271 TAAAGAACACACACTCTGAACCCAAACAAACAAAGATGCTTTGGCTAAGCAGCTGGGATT 330

OY 512 gacggaagaagaagtggaagtgtgttcacaaacagaagaagctag 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331 GCGACCCGAGCAAGTTGAGTTGTTCCAAAACAGAAAGGCGGAG 376
```

Search completed: June 19, 2001, 17:00:51
Job time: 22534 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2001, 17:00:28 ; Search time 7917.7 Seconds
(without alignments)
1212.991 Million cell updates/sec

Title: US-09-394-519-43

Perfect score: 1016
Sequence: 1 tgggaagccacaataacccc.....aaaaaaaaaaaaaaaa 1016

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	474.6	46.7	486	104	AI995124	AI995124 701502113
2	460.6	45.3	587	158	H76020	H76020 17725 Lambda
3	236.6	23.3	382	188	T22223	T22223 4231 Lambda
4	173.4	17.1	490	167	BE462282	BE462282 EST324546
5	172.6	17.0	418	118	AM626121	AM626121 EST320028
6	165.8	16.3	367	151	BF635572	BF635572 NF10401D
7	157	15.5	475	112	AM186273	AM186273 se65f12.y
8	156	15.4	389	16	AI100673	AI100673 35048 Lam
9	152.6	15.0	335	167	BE462229	BE462229 EST324493
10	139.6	13.7	685	75	AM109014	AM109014 gate00020
11	139.6	13.7	686	162	BE054276	BE054276 GA_Ea000
12	137.2	13.5	621	113	AM255172	AM255172 ML160 pep
13	137	13.5	447	174	BM134230	BM134230 EST467122
14	135.2	13.3	355	114	AM278294	AM278294 s142c03.y
15	134.4	13.2	905	153	BG445213	BG445213 GA_Ea002
16	134.4	13.2	913	76	AM730706	AM730706 GA_Ea002
17	133	13.2	431	148	BF424736	BF424736 su37h01.y
18	133	13.1	453	175	BE237350	BE237350 sab12c12.
19	131.4	12.9	857	146	BE269948	BE269948 GA_Eb000
20	131.2	12.9	945	153	BG446185	BG446185 GA_Ea003
21	131.2	12.9	953	77	AM731502	AM731502 GA_Ea003
22	130.8	12.9	585	151	BF650596	BF650596 NF095A01E
23	130.2	12.8	738	75	AM109501	AM109501 gate00047
24	130.2	12.8	742	162	BE054957	BE054957 GA_Ea000
25	130	12.8	633	153	BG448527	BG448527 NF036F04R
26	129.4	12.7	512	110	AM030476	AM030476 EST73731
27	129.4	12.7	626	110	AM032652	AM032652 EST76211
28	129.2	12.7	721	174	BM123386	BM123386 EST469032
29	129	12.7	660	155	BG592600	BG592600 EST491278
30	128.6	12.7	899	153	BG445739	BG445739 GA_Ea002
31	128.6	12.7	907	76	AM728013	AM728013 GA_Ea002
32	128.2	12.6	579	110	AM032348	AM032348 EST75802
33	128.2	12.6	691	110	AM032347	AM032347 EST75801
34	128	12.6	656	119	AM686107	AM686107 NF038C07N
35	127.8	12.6	671	153	BG447824	BG447824 NF103E12E
36	127.6	12.6	507	164	BE203702	BE203702 EST396378
37	127.6	12.6	528	166	BE352650	BE352650 EST016 So
38	127.6	12.6	605	21	AI489878	AI489878 EST24817
39	127.6	12.6	630	166	BE320090	BE320090 NF023F09R
40	127.6	12.6	655	151	BF648811	BF648811 NF050A07E
41	127.6	12.6	662	151	BF645356	BF645356 NF039H12E
42	127.6	12.6	663	151	BF648004	BF648004 NF041C03E
43	127.6	12.6	674	164	BE203901	BE203901 EST396577
44	127.6	12.6	681	151	BF644773	BF644773 NF020H12E
45	127.6	12.6	806	123	AM981167	AM981167 EST392361

ALIGNMENTS

RESULT	1	ALIGNMENTS
LOCUS	AI995124	486 bp mRNA
DEFINITION	701502113 A. thaliana, Ohio State clone set Arabidopsis thaliana	EST 08-SEP-1999
ACCESSION	AI995124	CDNA clone 701502113, mRNA sequence.
VERSION	AI995124.1	GI:5842029
KEYWORDS	EST	
SOURCE	thale cress.	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
AUTHORS	1 (bases 1 to 486)	
	Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R.	

TITLE

Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrrega, A., Murry, L., Turner, C., Krikorian, S., Elder, L., and Hanson, D.

JOURNAL

Arabidopsis thaliana Gene Expression Microarray

COMMENT

Unpublished (1999)

CONTACT

Contact: David Smoller, Ph.D.

SYSTEMS

Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.

ADDRESS

4633 World Parkway Circle, St. Louis, MO 63134, USA

TEL

Tel: 877-577-2733

FAX

Fax: 314-427-3324

EMAIL

Email: service@genomesystems.com.

FEATURES

Location/Qualifiers

1. 486

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/clone="701502113"

/clone_1lb="A. thaliana, Ohio State clone set"

/note="CDNA library was made from selected clones from the Arabidopsis thaliana Ohio State clone set."

BASE COUNT

148 a 83 c 111 g 144 t

ORIGIN

Query Match 46.7%; Score 474.6; DB 104; Length 486;
Best Local Similarity 99.2%; Pred. No. 1,9e-86;
Matches 477; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	1	ttggaagcccaataaccctatctcgtgcttttaaaagttagaataatcc	60
DB	1	ttggaagcccaataaccctatctcgtgcttttaaaagttagaataatcc	60
QY	61	gataaaactttatataatcttcttgatcagagaggtctgccaaggttga	120
DB	61	gataaaactttatataatcttcttgatcagagaggtctgccaaggttga	120
QY	121	ggaagagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc	180
DB	121	ggaagagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc	180
QY	181	gagaagcgaatgcatcagtcagtcagtcagtcagtcagtcagtcagtcagtc	240
DB	181	gagaagcgaatgcatcagtcagtcagtcagtcagtcagtcagtcagtcagtc	240
QY	241	gtgacacgaatgcatcagtcagtcagtcagtcagtcagtcagtcagtcagtc	300
DB	241	gtgacacgaatgcatcagtcagtcagtcagtcagtcagtcagtcagtcagtc	300
QY	301	atgaagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc	360
DB	301	atgaagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc	360
QY	361	ctggtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc	420
DB	361	ctggtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc	420
QY	421	gttaaaacatgagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc	480
DB	421	gttaaaacatgagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc	480
QY	481	c 481	
DB	481	c 481	

SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis.
1 (bases 1 to 382)

REFERENCE
AUTHORS
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Ralkehl,N., Somerville,S., Thomasow,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)

TITLE
JOURNAL
MEDLINE
95148729
COMMENT
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@lhm.cl.msu.edu
Seq primer: T7 dye primer.

FEATURES
source
1..387
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="193M157"
/clone_1lb="Lambda-PRU2"
/note="Vector: Lambda Zip-Lox; Site_1: Sal; Site_2: Not;
Lambda PRU2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."

BASE COUNT
165 a 106 c 135 g 167 t 14 others

ORIGIN
Query Match 45.3%; Score 460.6; DB 158; Length 587;
Best Local Similarity 94.2%; Pred. No. 1.3e-83;
Matches 505; Conservative 0; Mismatches 28; Indels 3; Gaps 3;

QY 2 gggagccacaataacccccctatctctggcctttttaaaagttagaataatccg 61
DB 3 GGGNNCCACAAATMACCCCTATTCCTCGCCCTTTTAAAAAGTTTAAATATCCG 62

OY 62 ataaatctctatataatatttcttgatgcatgaggggtgcgcaaggggttag 121
DB 63 ATAAATCTCTTTATTTATTTTNTTGGTCATGAGAGGGTTCGCCAAAGGTTGAG 122

OY 122 gaaagtgatgactgctgaaagaagatagctctctgaggtatgataagatag 181
DB 123 GAAAGTGATGACGTGCAAGAGATAGTCTGTGAGGCTATGTATGATATAGTATGG 182

OY 182 agaaagcaatgcatcaagttctcttgagagtggtggtcgaatcgatcgagaaggttg 241
DB 183 AGAAGCAATGCGATCAAGTTCCTTTGAGAGCTGGCTTAAATCGATGCAAGAGTTG 242

OY 242 tagacaagaatgctgaactattgagccaagtatcaagaagggagaactgacaatg 301
DB 243 TAGACATAAGATGTTTAACTATTGGAAGCCAAATATCAAGAGGAGAGACTTATGCAATGA 302

OY 302 tgaagttagctctctctctctcctcataaagctctcagaanaata-gtctgcttgatg 360
DB 303 TGAAGTTGATCTTCTCTCTGCTGCTCATTAAGCTTATGAGGAATAGGCTGCTTATG 362

OY 361 ctggtctgactgctggtgagccgctaatgagtcaaaaatta-ctggaacaccatcgt 419

DB 363 CTGTCGATTGCTGTCGCGACCGCTAATGATGCAAAAATTACTGGAACACCATCTG 422

OY 420 agtaaaaaacatgagctctcggttgaagtctaataataaaaagaacatttttc 479
DB 423 AGTAAAAACATGAGNGCTTGCTGTTGATGAGCTTAATATAAAGGAAACCATATTTC 482

OY 480 cctctcacaca-ccgctcacaanaacgcgtgtttttaaagctcgacatcgtcct 534
DB 483 CCTCCTNCACANCCGCGNCAAAAANCGGTGTTTAAAGCTTGACCTGCGTCCCT 538

RESULT 3.
LOCUS T22223 382 bp mRNA EST 06-NOV-1997
DEFINITION 4231 Lambda-PRU2 Arabidopsis thaliana cDNA clone 99P97, mRNA
sequence.
ACCESSION T22223
VERSION T22223.1 GI:2596878
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis.
1 (bases 1 to 382)
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Ralkehl,N., Somerville,S., Thomasow,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)

JOURNAL
MEDLINE
95148729
COMMENT
On Nov 6, 1997 this sequence version replaced gi:932253.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRU, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@lhm.cl.msu.edu
Seq primer: T7 dye primer.

FEATURES
source
1..382
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="99P97"
/clone_1lb="Lambda-PRU2"
/note="Vector: Lambda Zip-Lox; Site_1: Sal; Site_2: Not;
Lambda PRU2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."

BASE COUNT
104 a 64 c 89 g 110 t 15 others

ORIGIN
Query Match 23.3%; Score 236.6; DB 188; Length 382;
Best Local Similarity 82.7%; Pred. No. 3.2e-38;
Matches 286; Conservative 0; Mismatches 57; Indels 3; Gaps 2;

OY 71 tttaataaattttcttggtcatalggaaggttcgccaagggctgagaaggtgc 130
DB 23 TATTTAACGATCATCTTGTTCATGAGAGGGTTCGCCAAAGGCTCGCAAAAGTGC 82

OY 131 atgactctgaaagaagatagctctcttgaggtatgtatgataagtagagaagcaa 190

```

Query Match      17.13: Score 173.4; DB:167; Length 490;
Best Local Similarity 70.6%; Pred. No. 2.1e-25;
Matches 231; Conservative 0; Mismatches 96; Indels 0; Gaps
0
0y 101 gggctgcaccaagggtttgaaagatgcatgactgctgtaagaagatagctctctgag 166
17 ggtgcacacgtttggcagattagcaaaagttctatggacacgtcaacaaagaagttctcttttaag 76

```

	Query Match	Similarity	Score	DB	Length
Db	17	97.78%	172.6	DB	418
	Matches	226	Conservative	0	Mismatches 89; Indels 0; Gaps
Qy	101	gggttcgtccaaagggtttgagaaagtgatgcatgactgctgtaagaagatagtcctcttgag	160		
	17	gggttcacatcgtttggagattagaaagggttcactgacgtgacacaaagagttctcttttaag	76		

Qy	183	gaagcgcaatgagcgtcaagttcttgagagcgtgagcctaactcgatgcagaaaggtgtc	242
Db	62	GAAGGAAAATGGCAATTAGTTCCTAAACGAACAGGTTTGATAGGTGGTCCGAAAAGTTGT	121
Qy	243	agactaaagatggtctggaactattgaagccaaagtaaccaagagagaaagacttaagcaatg	302
Db	122	AGATTGAGGTGGTTGAATTATTATTTAAACCCCGCATCAACAGAGAAATTTTCGCCGAGAT	181
Qy	303	gaattgatctctctcttcgtccctcaaaagctctaggaataagtggtgtcctgattgtc	362
Db	182	GAAATTGATATGATTCTPAAGGTTACACAAACTTCTAGGAAACAAATGCTTTTGATTGCT	241
Qy	363	ggtcgtatgctctgctgcagccgccaatgactgcaaaaattactggaaccaccactgagt	422
Db	242	GCAAGGCTCCGGGTAGGACAGCTAATGATGTGAAANATTATTGGCACCAAAATTTGCCG	301
Qy	423	aaaaaa 428	
Db	302	AAAAAA 307	

LOCUS	475 bp	mRNA	EST	18-APR-2000	
DEFINITION	AM186273				
	gag5515.y1 Gm-cj019 glycine max CDNA clone				GENOME SYSTEMS CLONE ID
	Gm-cj019-1896 5' similar to TR:Q40920 Q40920 MYB-LIKE				
	TRANSCRIPTIONAL FACTOR MBFL. ; , mRNA sequence.				
ACCESSION	AM186273				
VERSION	AM186273.1				
KEYWORDS	GI:6455590				

ORGANISM
Glycine max
Euparvota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustersids I; Fabales; Fabaceae; Papilionoideae; Glycine.
1 (bases 1 to 475)
Shenmaker, R., Kelm, P., Vodka, L., Expelding, J., Corryell, V., Khana
A., Bolla, B., Marre, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck
R., Rittner, E., Kohn, S., Shinn, T., Jackson, V., Cardenas, M., McCann

TITLE
JOURNAL
COMMENT

R, Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project:
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Inc. 4633 WC

maple street is a 400000 bp long contig. Clones of approximately 1000 bp size
Parkway Circle St. Louis, Missouri 63134 for further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Insert Length: 957 Std Error: 0.00
Seq primer: -40RP from Gldco
High quality sequence stop: 443.
location/qualifiers
1. .475
/organism="Glycine max"
/db_xref="taxon:3847"

Greenhouse grow

```

//clone_id="G1019
//tissue_type="Immature seed coats of greenhouse grown
plants"
/lab_host="DR10B (Gibco BRL)"
/note="Vector: pSPOR1 (Life Technologies); Site_1: Not I;
Site_2: Sal I; This cDNA library was constructed from mRNA
isolated from immature seed coats (200-300 mgs) of
greenhouse grown plants. The library was prepared using

```


is being submitted in the sense orientation
Seq primer: M13_Universal.

Seq primer: M13_universal.

FEATURES

Source

Source

1.389

/organism="Arabidopsis thaliana"

```
/strain="var columbia"
```

```
/db_xref="taxon:3702"
/c)one="qqpyp"
```

```
/clone="yyppyXP"
```

```
/clone_id= lambda -fklz
/note="Vector: lambda 2i
```

Lambda PRL2 is a cDNA library derived from equal

quantities of 4 pools of mRNA. The mRNA sources were 1) 7

day germinated etiolated seedlings; 2) tissue culture

grown roots; 3) staged plants half with 24 hour light

cycle, half on 16 hr light, 8 hour dark-rosettes; 4)

LOCUS BE054276 686 bp mRNA EST 07-MAR-2001
DEFINITION GA_Ea002018f Gossypium arboreum 7-10 dpa fiber library Gossypium
ARBOREUM cDNA clone GA_Ea002018f, mRNA sequence.
ACCESSION BE054276
VERSION BE054276.2 GI:13244153
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Malvales; Malvaceae; Gossypium.
REFERENCE 1 (bases 1 to 686)
Wing, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
, D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
On Jun 8, 2000 this sequence version replaced gi:8381332.
JOURNAL Contact: Wing RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATAGACTACTATAGCG
High quality sequence stop: 559.
FEATURES
source
1..686
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea002018f"
/clone_1lb="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 228 a 142 c 146 g 170 t
ORIGIN
Query Match 13.7%; Score 139.6; DB 162; Length 686;
Best Local Similarity 65.3%; Pred. No. 1.5e-18;
Matches 205; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
QY 113 aggggtgaaggaaggtgcatgactgctgaagaagatagctcttggagctatgtatga 172
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 155 AGGGTGAAGAAAGGTCCATGACCCGAGAAAGATCAAAAGCTTACCTTACATTTGA 214
QY 173 taagtgtggaaggaaggaagcatcaagtctcttggagctggtgctaatgacag 232
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 215 ACAACATGGCCATGAGAGCTGGCGTCTTAAAGCTGGGCTTCAAAAGATGTGG 274
QY 233 aaagagttgagactaagatggttgaactattgaagccaagatcaagaaggaagact 292
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 275 AAAGAGTTGAGAGCTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 334
QY 293 tagcaatgataagatgatactcttcttgcctcctcaataagcttcaagaatagtgctc 352
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 335 CAGTTTACAAAGAAAGAACGACATTAATTAATTAATTAATTAATTAATTAATTAATTA 394
QY 353 ctgattgtcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 412
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 395 TGGCATTAGCTACTTATTTGCCGAAAGAACAGACAAATGATCAAGACTTCTGGAACAC 454
QY 413 ccactgtagtaaaa 426
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 455 ACATCTTAAGAAAAA 468
RESULT 12

AM255172
LOCUS AM255172 621 bp mRNA EST 23-AUG-2000
DEFINITION M160 peppermint glandular trichome Mentha x piperita cDNA, mRNA
sequence.
ACCESSION AM255172
VERSION AM255172.1 GI:7244424
KEYWORDS EST.
SOURCE peppermint.
ORGANISM Mentha x piperita
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Lamiales; Lamiales; Mentha.
REFERENCE 1 (bases 1 to 621)
Lange, B.M., Wildung, M.R., Stauber, E.J., Sanchez, C., Pouchnik, D. and
Croiseau, R.
Probing essential oil biosynthesis and secretion by functional
evaluation of expressed sequence tags from mint glandular trichomes
Proc. Natl. Acad. Sci. U.S.A. 97 (6), 2934-2939 (2000)
JOURNAL MEDLINE
20183992
COMMENT Contact: Lange, B.M.
Institute of Biological Chemistry/Washington State University
Pullman, WA
Email: Lange-m@mail.wsu.edu.
FEATURES
source
1..621
/organism="Mentha x piperita"
/cultivar="Black Mitcham"
/db_xref="taxon:34256"
/clone_1lb="peppermint glandular trichome"
/tissue_type="pelate glandular trichomes"
/cell_type="secretory"
/note="Vector: lambda ZAPit"
BASE COUNT 204 a 130 c 163 g 124 t
ORIGIN
Query Match 13.5%; Score 137.2; DB 113; Length 621;
Best Local Similarity 63.8%; Pred. No. 4.5e-18;
Matches 208; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 113 aggggtgaaggaaggtgcatgactgctgaagaagatagctcttggagctatgtatga 172
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 60 AGGGCTGAAGAGAGGAGTGGAGTCTGTAAGAGACAAACTAATAATTAATTTCA 119
QY 173 taagtgtggaaggaaggaagcatcaagtctcttggagctggtgctaatgacag 232
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 120 GGAATAATGCTGAAGCTGCTGGAGATCATTTGCCAAGAAATGCAAGTTTACTTATGATGTGG 179
QY 233 aaagagttgagactaagatggttgaactattgaagccaagatcaagaaggaagact 292
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 180 AAAGAGTTGAGAGCTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 239
QY 293 tagcaatgataagatgatactcttcttgcctcctcaataagcttcaagaatagtgctc 352
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 240 TTCTTCAAGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 299
QY 353 ctgattgtcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 412
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 300 CCTGATGCCCGCGACTTGGCGGTAGAACAGACAAATGAATTAATTAATTAATTAATTAATTA 359
QY 413 ccactgtagtaaaaacatgagctt 438
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 360 CCATTGTAGCAGAAATTTTCACGCTT 385
RESULT 13
BG134230
LOCUS BG134230 447 bp mRNA EST 31-JAN-2001
DEFINITION EST467122 tomato crown gall Lycopersicon esculentum cDNA clone
CTOE15D4 5' sequence, mRNA sequence.
ACCESSION BG134230
VERSION BG134230.1 GI:12634418
KEYWORDS EST.

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** Small print

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)
